Genetic and epigenetic regulation in spermatogenesis in mice; lessons from *Alkbh1* and *Tzfp* mutants

Kari Furu

Thesis for the degree of Philosophiae Doctor (PhD)

Centre for Molecular Biology and Neuroscience (CMBN)

Department of Microbiology

Oslo University Hospital HF, Rikshospitalet

University of Oslo, Norway

2012

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1427

ISBN 978-82-8264-216-3

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Acknowledgements

The work presented in this thesis has been performed at the Institute of Medical Microbiology at Oslo University Hospital, Rikshospitalet, in 2009-2012. I was lucky to get employed by Professor Arne Klungland after I had finished my Master's degree in Trondheim, and after working in his lab as a technician for one year, I started my Ph.D. The work is funded by the Norwegian Cancer Society.

I want to express my gratitude to my supervisor Professor Arne Klungland and my co-supervisor Professor Magnar Bjørås. I am grateful for the opportunity to work in such a great work environment with so many gifted and supportive people. Thankfully, I've been given the freedom to develop my own project and to test many of the hypotheses I've come up with, many of which have lead nowhere, but has thought me scientific thinking and the art of being patient!

Furthermore, I wish to thank my colleagues at the Institute. Science can be painfully frustrating from time to time, and I don't think I could have completed this work if it wasn't for the great gang working together with me on my projects and sharing the cramped office space. I appreciate the many laughs, coffee breaks and scientific discussions, as well as the occasional "Thank God It's Friday"-parties. I have really enjoyed working with you all! I am forever grateful to the contributors who have made this work possible; co-workers, technicians, professors, students, post-docs, the animal facility, statisticians etc. Thanks also to the co-authors of the papers listed in this thesis. In addition, I would like to highly acknowledge John-Arne Dahl for useful scientific discussions and for critically reading this thesis.

I want to express my gratitude to my family and friends who have endured my frustrated complaints when things haven't gone my way in the lab. Finally, a big and loving thank you to my wonderfully supportive Anders.

Contents

List of Papers

This thesis is based on the following papers, which will be referred to by their Roman numerals:

I Nordstrand L.M., Svärd J.*, Larsen E.*, Nilsen A.*, Ougland R.*, **Furu K.***, Lien GF, Rognes T, Namekawa SH, Lee JT, Klungland A. "Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects" (* contributed equally)

PLoS One. 2010 Nov 3;5(11):e13827

II Nordstrand L.M.*, **Furu K.***, Paulsen J., Rognes T., Klungland A. "Alkbh1 and Tzfp Repress a Non-Repeat piRNA Cluster in Pachytene Spermatocytes" (* contributed equally)

Nucleic Acids Research. 2012, in press

III **Furu K.**, Klungland A. "Tzfp represses the Androgen Receptor in mouse testis"

Manuscript (2012)

Abbreviations

1meA	1-methyladenine
2OG	2-oxoglutarate
3meC	3-methylcytosine
5meC	5-methylcytocine
5hmeC	5-hydroxymethylcytocine
6meA	6-methyladenine
α-KG	alpha-ketoglutarate
ALKBH	AlkB homolog
AR	Androgen Receptor
BMP	Bone Morphogenetic Protein
BTB/POZ	Broad complex, Tramtrack, Bric a Brac or Pox virus and Zinc finger
CB	chromatoid body
DNMT	DNA methyltransferase
dpc	days post coitum
dpp	days post partum
Drosophila	*Drosophila melanogaster* (fruit fly)
DSB	double stranded break
E. coli	*Escherichia coli*
FA	Fanconi Anemia
FANCC	Fanconi Anemia complementation group C
Fe	Iron
JmjC	jumonji domain-containing
HAT	Histone Acetyl Transferase
HDAC	Histone deacetylase
Het	heterozygous
HP1	Heterochromatin Protein 1
IAP	Intracisternal A Particle
ICM	intermitochondrial cement
L1	Long Interspersed Element 1
KO	knockout
miRNA	microRNA
MSCI	meiotic sex chromosome inactivation
MSUC	meiotic silencing of unsynapsed chromosomes
mRNA	messenger RNA
N-CoR/SMRT	nuclear receptor, co-repressor and silencing mediator for retinoid and thyroid hormone receptors complex
ncRNA	non-coding RNA
PGC	primordial germ cells
Plzf	Promyelocytic leukemia zinc-finger
piRNA	PIWI-interacting RNA
PTM	post-translational modification
RNAi	RNA interference
S. pombe	*Schizosaccharomyces pombe*
SAM	S-adenosyl methionine
SC	Sertoli cells
SUMOylation	small ubiquitin-like modification
tbs	Tzfp binding sequence
TE	transposable elements
TRD	transmission ratio distortion
Tzfp	Testis zink finger protein
WT	wild-type
Zn	zinc

Glossary

BTB/POZ domain	protein domain that mediates homomeric, oligomeric and heteromeric dimerisation
C_2H_2 zinc fingers	a commonly found protein domain often found in regulatory proteins in mammals. Have DNA-binding properties that depend on the presence of zinc ions
chromatin	mass of genetic material composed of DNA and proteins that condense to form chromosomes during eukaryotic cell division
dioxygenase	an enzyme that catalyzes the insertion of an oxygen molecule into an organic substrate
ectoderm	one of the three primary germ cell layers in the early embryo. Differentiates to form the nervous system, tooth enamel and the epidermis
euchromatin	uncondensed, active chromatin
GATA	a family of transcription factors characterized by their ability to bind to the DNA sequence "GATA" Consists of at least 6 members (GATA-1-6) that play crucial roles in the development and differentiation of all eukaryotic organisms
heterochromatin	condensed, inactive chromatin
histones	small, basic proteins interacting with DNA to form nucleosomes and chromatin that is essential for DNA compaction and epigenetic regulation
incomplete penetrance	aphenomenon in which some individuals fail to express a trait, even though they carry the allele
piRNAs	class of small non-coding RNAs 24-31 nt in length that form complexes with PIWI proteins. In mammals, piRNAs are testis specific and have been linked to silencing of retrotransposons
piRNA cluster	genomic region encoding a primary piRNA precursor. May contain as few as ten or up to thousands of piRNAs. Often located in intergenic regions or in transposon rich areas
pachynema	the stage in prophase I of meiosis where synapsis of homologous chromosomes is completed and crossing over occurs
primary piRNA precursor	a single RNA molecule transcribed from a piRNA cluster resulting in the production of primary piRNAs
prophase I of meiosis	the first prophase of meiotic cell division. Involves chromosome condensation and synapsis, crossing-over, the disappearance of the nucleolus, meiotic spindle formation, and disappearance of the nuclear envelope. Is devided into leptonema, zygonema, pachynema and diplonema
protamines	small, nuclear proteins that replace histones during the elongated spermatid stage of spermiogenesis. believed to be essential for the dense DNA condensation that occurs in the maturing spermatozoa
spermatogenesis	the development of germ cells in male organisms through meiotic division of primary spermatocytes

spermiogenesis	the final stage of spermatogenesis, which involves the maturation of spermatids into mature, motile spermatozoa
synapsis	the intimate pairing of the parental chromosomes during meiosis
transposon	a type of mobile genetic element that can change its relative position (self-transpose) within the genome of a single cell
trophoblast	cells forming the outer layer of a blastocyst, which will develop into a large part of the placenta

Summary

Epigenetics describe heritable changes in gene expression and cellular phenotypes caused by mechanisms other than changes in the underlying DNA sequence. This thesis describes biological properties of the Alkbh1 and Tzfp proteins, and discusses their role in epigenetic and developmental processes. Based on work performed on transgenic mice, we provide evidence for a role in spermatogenesis for these two proteins.

Alkbh1 is a mouse homolog to the *E. coli* DNA repair enzyme AlkB, which directly repairs DNA lesions such as 1meA and 3meC. These proteins belong to a family of dioxygenases that utilizes 2-oxoglutarate as a substrate and Fe^{2+} as a co-substrate to remove methyl groups from target biological molecules. Tzfp belongs to the BTB/POZ zinc finger family, a group of proteins that play key roles in many biological processes both in vertebrates and lower organisms through regulating the transcriptional activity of downstream genes. Tzfp is highly expressed in testis and is shown to work as a transcription repressor.

To investigate the role of Alkbh1 and Tzfp, we generated two novel mutant mice strains; *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$. In paper I, we describe the *Alkbh1* deficient mouse and the phenotypes observed. Lack of Alkbh1 leads to reduced viability during embryonic development and an incompletely penetrant phenotype of recurrent unilateral eye defects and craniofacial malformations. Additionally, the sex-ratio is considerably skewed against female offspring. *Alkbh1* is highly expressed in testes, and apoptotic spermatids were revealed in 5-10% of the tubules in *Alkbh1*$^{-/-}$ testes, suggesting a role in spermatogenesis. Combined, these data indicate that Alkbh1 is involved in differentiation, particularly during spermatogenesis and embryonic development.

In Paper II we identify a role for Alkbh1 and Tzfp in the regulation of piRNA in mouse testes. This links Alkbh1 and Tzfp to the mouse PIWI-piRNA pathway, a testis specific set of regulatory mechanisms involved in diverse biological and molecular functions including spermatogenesis and transposon silencing. Using *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ mice, we show a remarkable regulation of a single piRNA cluster, which we call 1082B. Cluster 1082B is a unidirectional, non-repeat associated cluster, encoding a precursor transcript containing 788 individual piRNAs. Finally, we propose a role for the 1082B-derived piRNAs in transposon control, as the remarkable upregulation of the 1082B-derived piRNAs was followed by a downregulation of LINE1 and IAP transcripts in testes from both *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ mice.

In paper III I describe the *Tzfp*$^{GTi/GTi}$ mouse. Mice lacking Tzfp are viable and fertile with no obvious phenotype. A sex ratio distortion was observed upon breeding heterozygous animals,

skewed against male offspring. The Tzfp gene was found to be highly expressed during the pachytene stage of meiosis in male gametogenesis, indicating a role in the formation of spermatozoa. Several genes were found to be upregulated in mutant testes, confirming previous data suggesting that Tzfp is a transcription repressor. Additionally, Androgen Receptor signaling was found to be highly increased in Sertoli cells in $Tzfp^{GTi/GTi}$ testes.

1 Introduction

1.1 Epigenetics

Epigenetics has been defined as the study of heritable changes in gene expression or cellular phenotype caused by mechanisms other than changes in the underlying DNA sequence (Berger *et al.*, 2009). All nucleated cells in the body contain the same DNA sequence (with a few exceptions), though they vary greatly in terms of function, size and the ability to regenerate. To achieve the formation of different cell types within an organism without altering the genome, cells use epigenetic mechanisms to modulate gene expression (Figure 1) (Morgan *et al.*, 2005; Bird, 2002; Turner, 2007; Ringrose and Paro, 2004).

Epigenetics consists of two main components, both of which play a role in regulating gene expression. The first is the addition of methyl groups to the backbone of DNA. The second is a variety of covalent posttranslational modifications added to the histone proteins, affecting how tightly DNA is wound around the histone cores. These chemical modifications in our genome make up our epigenome and are affected by several factors, such as development *in utero* and exposure to environmental chemicals, drugs and dietary components (Mittal et al., 2003; Marsit et al., 2006). The epigenome also changes with age (Richardson, 2003b). Incorrect epigenetic regulation may result in diseases such as cancer, autoimmune disease, or diabetes (Chen et al., 1998; Richardson, 2003a; Maier and Olek, 2002).

1.1.1 DNA methylation

DNA methylation refers to the addition of a methyl group to the 5th carbon of cytosine by a DNA methyltransferase using S-adenosyl methionine (SAM) as a substrate, generating 5-methylcytosine (5meC) (Figure 1). In adult somatic tissues, DNA methylation typically occurs in a CpG dinucleotide context, whereas non-CpG methylation is prevalent in embryonic stem cells (the "p" in CpG notation refers to the phosphodiester bond between the cytosine and the guanine) (Lister *et al.*, 2009). CpGs are often grouped in clusters called CpG islands, which are present in the 5' regulatory regions of many genes. In general, CpG island methylation is associated with gene silencing (Bourc'his and Bestor, 2004; Kaneda *et al.*, 2004; Rhee *et al.*, 2002; Larsen *et al.*, 1992; Bird *et al.*, 1985). Most CpG islands are unmethylated, including those located at the 5' ends of housekeeping genes (Straussman *et al.*, 2009). DNA methylation also occurs at repetitive sequences, gene body and at CpG island shores.

DNA methylation is essential for normal development (Li *et al.*, 1992) and is associated with a number of key processes including genomic imprinting (Li *et al.*, 1993), X-chromosome

inactivation (Panning and Jaenisch, 1998), suppression of repetitive elements (Walsh *et al.*, 1998), carcinogenesis (Jones, 1996), and cell aging (Ahuja *et al.*, 1998). Three DNA methyltransferases (DNMTs) have been identified in mammals; DNMT1, DNMT3A and DNMT3B (Bourc'his and Bestor, 2004; Kaneda *et al.*, 2004). They are responsible for the generation of new DNA methylation patterns, as well as maintaining the existing pattern upon cell division. DNMT3L is a protein closely related to DNMT3A and DNMT3B, and is critical for DNA methylation, but appears to be enzymatically inactive on its own.

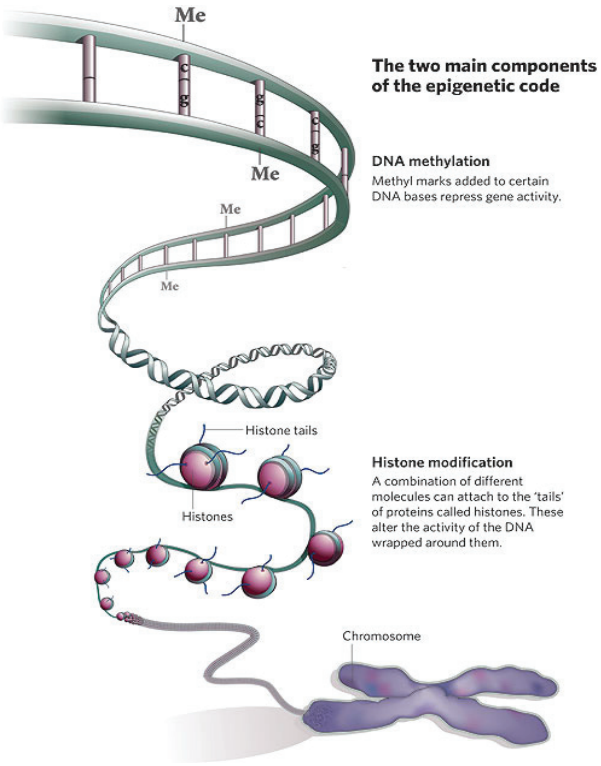


Figure 1. Chromatin and epigenetics. The two main components of epigenetics; DNA methylation and posttranscriptional histone modification, are indicated. The addition of methyl groups to cytosine bases, forming 5meC, repress gene activity, whereas addition and removal of chemical groups to the amino acid residues of histone tails regulates DNA activity by affecting DNA condensation (With permission from Qiu, 2006).

DNA methylation patterns are traditionally thought to be relatively stable and have an important role in long-term silencing, creating methylation patterns that persist upon cell division. Recent studies have, however, revealed that DNA methylation is more dynamic than originally thought, with methylation and active demethylation occurring during specific stages of development, such as in the paternal pronucleus of a zygote right after fertilization (Oswald *et al.*, 2000), and in the primordial germ cells (PGCs) of developing embryos

(Hajkova *et al.*, 2002). Demethylation of 5meC residues in DNA is proposed to involve various mechanisms, such as by the action of 2-oxoglutarate- and Fe(II)-dependent oxygenases of the by the ten-eleven translocation (TET) enzymes that creates 5-hydroxymethylcytocine (5hmeC) (Tahiliani *et al.*, 2009). 5hmC may be an intermediate that is further processed, ultimately restoring the cytosine base (Boorstein *et al.*, 2001; Ito *et al.*, 2010) or, alternatively, 5hmC may itself be a functional modification (Kriaucionis and Heintz, 2009). Recent studies favour a role for 5hmC as an epigenetic mark distinct from 5mC (Doege *et al.*, 2012). 5hmeC has been reported in Purkinje cells, granule cells and in mouse embryonic stem cells, but seems to be very low or absent in cancer cell lines (Booth *et al.*, 2012; Kriaucionis and Heintz, 2009).

1.1.2 Histone modifications

Histones are proteins found in eukaryotic cell nuclei in association with DNA. They package and order the DNA into structural units called nucleosomes. Together with DNA, they make up the main component of chromatin and act as spools around which DNA winds. The core histones H2A, H2B, H3 and H4 form a nucleosome octamer with a 147-bp segment of DNA wrapped around it. Neighboring nucleosomes are separated by approximately 50 bp of free DNA. The histones are subject to a large number of post-translational modifications that affect chromatin compaction (Figure 1) (Kouzarides, 2007); (Wolffe and Pruss, 1996). Histone modifications play a large role in transcriptional regulation, but are also found to have important roles in DNA replication and repair (Botuyan *et al.*, 2006), alternative splicing (Luco *et al.*, 2010) and chromosome condensation (Bartke *et al.*, 2010; Shogren-Knaak *et al.*, 2006).

Most histone modifications have been found to be reversible, creating a flexible epigenetic pattern that can be altered to allow changes in gene expression in the cell. Such changes are brought about by enzymes catalyzing the formation and removal of the modifications. There are at least eight distinct types of chemical modifications found on histones; acetylation (Sterner and Berger, 2000), methylation (Zhang and Reinberg, 2001), phosphorylation (Nowak and Corces, 2004), ubiquitylation (Shilatifard, 2006), SUMOylation (Nathan *et al.*, 2006), ADP ribosylation (Hassa *et al.*, 2006), deimination (Cuthbert *et al.*, 2004) and proline isomerization (Nelson *et al.*, 2006).

Some histone modifications cause the chromatin to form a compact, inactive structure, called heterochromatin. Constitutively condensed chromatin mostly contains repetitive DNA, like the centromeres and telomeres. In other parts of the genome, heterochromatin is associated with inactive, untranscribed DNA. These areas can lose their condensed structure and become transcriptionally active. Other histone modifications cause the chromatin to form an open, euchromatic state, which is associated with actively transcribed genes. The two different

forms are characterized by a certain subset of histone marks (Bannister and Kouzarides, 2011; Kouzarides, 2007). Heterochromatin is often associated with di- and tri-methylated H3K9 and H3K27, whereas euchromatin is characterized by high levels of tri-methylated H3K4 in promoter regions (Szutorisz *et al.*, 2005; Karlic *et al.*, 2010) and acetylation of lysine residues on histone H3 and H4 (Roth and Allis, 1996). In addition, H3K36me3 and methylated H3K4 are highly enriched in actively transcribed regions (Bannister *et al.*, 2005; Pokholok *et al.*, 2005).

Acetylation marks of lysine residues are highly dynamic and are regulated by the opposing action of two families of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs), whereas histone methylation, which normally occurs on lysine and arginine residues, are regulated by specific SAM-dependent methyltransferases (Bannister and Kouzarides, 2011; Kouzarides, 2007). Histone methylation has been considered a stable, static modification, but the last decade, a number of different demethylation mechanisms for both lysine and arginine have been identified, including demethylation by the jumonji-domain containing (JmjC) proteins (Tsukada *et al.*, 2006).

1.1.3 RNA-mediated phenomena

Cellular RNA, such as mRNA, tRNA, rRNA and snRNA have long been known to contain many structurally distinct post-transcriptional modifications (Cantara *et al.*, 2011; He, 2010), some of which may have regulatory roles similar to those of protein and DNA modifications (He, 2010). The study of RNA modifications and their effect on gene expression might be defined as the third level of epigenetic regulation, and is called RNA epigenetics. One example is the N(6)-methyladenosine (6meA) modification in mRNA which has been proposed to affect mRNA processing and export from nucleus to cytoplasm (Harper *et al.*, 1990). Recent findings suggest that 6meA plays a fundamental role in the regulation of gene expression (Dominissini *et al.*, 2012). The human obesity associated FTO protein has been found to be an RNA demethylase that oxidatively removes this modification in mRNA (Jia *et al.*, 2011).

Non-coding RNAs (ncRNAs) potentially play an active role in modulating gene transcription and epigenetic states. Recent studies indicate that the genome is extensively transcribed, producing a variety of short and long non-protein-coding RNAs (Bertone *et al.*, 2004), including transcripts from retrotransposons, which may play important roles in a variety of processes during differentiation and development (Mattick *et al.*, 2010). A role of an RNA interference (RNAi)-like mechanism in heterochromatin formation has been suggested in *S. pombe*, where transcripts from a transgene inserted into centromeric repeats gets processed into small interfering RNAs (siRNAs), resulting in efficient heterochromatic silencing (Halic and Moazed, 2010). Additionaly, in *Tetrahymena*, methylation of H3K27 is dependent on the

RNAi machinery (Liu *et al.*, 2007). In mammals, the *Xist* ncRNA is crucial for the silencing of one randomly chosen X chromosome in female animals (Sheardown *et al.*, 1997). In mice, the expression of the Argonaute-related protein Miwi2 and its associated pre-pachytene piRNAs (described below), is essential for transposon silencing through *de novo* DNA methylation in testes from about 15.5 days post coitum (dpc) to 3 days post partum (dpp) (Aravin *et al.*, 2008).

Another group of non-coding RNAs thought to play important epigenetic roles, are the microRNAs (miRNAs). miRNAs are RNA molecules (appriximately 23 nt long) found in eukaryotic cells, involved in post-transcriptional regulation (Brennecke *et al.*, 2005; Bartel, 2009). They bind target mRNAs, resulting in translational repression or target degradation and gene silencing. One example is the epigenetic silencing of the gene encoding the transcription factor NFI-A by miR-223 (Zardo *et al.*, 2012). Abnormal expression of this miRNA is associated with cancer (Fazi *et al.*, 2007).

1.1.4 Epigenetic crosstalk

DNA methylation may affect the transcription of genes by physically impeding the binding of transcriptional proteins to the gene. Secondly, methylated DNA may bind methyl-CpG-binding domain proteins (MBDs) which recruit HDACs and methyl transferases, affecting chromatin compaction through histone modifications. Recruited HDACs will remove the “activation marks” on H3 and H4, resulting in a compact, inactive chromatin. The addition of a methyl group can be recognized and bound by chromatin silencers such as HP1 (Bannister *et al.*, 2001), inducing an inactive state. This couples DNA methylation to transcriptional silencing through the modification of chromatin proteins (Prokhortchouk and Hendrich, 2002).

Other examples of such epigenetic cross-talk have also been identified. For example, the transcriptional repressor Polycomb Repressive Complex 2 (PRC2), a histone methyltransferase methylating H3K27, is negatively regulated by DNA methylation (Bartke *et al.*, 2010). This may enable PRC2 to associate preferentially with a specific chromatin state, enabling a more fine-tuned response to external stimuli. This provides evidence that proteins can monitor the methylation state of both histones and DNA in order to discriminate between distinct states of repressed chromatin.

1.1.5 Transposons and transposon silencing

Transposable elements (TEs), also known as "jumping genes," are DNA sequences that can move from one genomic location to another. Active transposons are potential insertional mutagens, capable of causing genome rearrangements (Burwinkel and Kilimann, 1998), exon skipping or alternative splicing (Takahara *et al.*, 1996). They are therefore silenced by the

host through epigenetic defense mechanisms such as DNA methylation (Walsh *et al.*, 1998), chromatin remodeling, and RNA interference (RNAi) (Carmell *et al.*, 2007; Kuramochi-Miyagawa *et al.*, 2008; Aravin *et al.*, 2008; Aravin *et al.*, 2007a; Aravin *et al.*, 2007b).

TEs are found in large numbers in the genome of many organisms and make up approximately 50% of the human genome (Lander *et al.*, 2001). Most transposons, like the Long terminal repeat (LTR)-containing Intracisternal A particle elements (IAPs), have accumulated mutations that render them unable to transpose (Dupressoir and Heidmann, 1996). In humans, the only type of actively transposing TEs is Long Interspersed Nucleotide Element 1 (LINE1 or L1), which make up an estimated 21% of the human genome (Lander *et al.*, 2001)

Mutations due to transposon transpositioning is a rare event, but has been recorded in many animals. Transposition of P elements in *Drosophila* may result in both germline and embryonic defects causing sterility through a phenomenon called hybrid dysgenesis (Rubin *et al.*, 1982). In mice, lack of Dnmt3l results in hypomethylated transposons in the testis, leading to abnormal and abundant expression of several transposon families (Bourc'his and Bestor, 2004). In humans, about 100 active L1 elements have been detected (Brouha *et al.*, 2003), and the degree of LINE methylation can be correlated with the degree of malignancy in certain cancers (Takai *et al.*, 2000; Roman-Gomez *et al.*, 2005). In addition, insertion of L1 into the *APC* tumor suppressor gene in colon cancer (Miki *et al.*, 1992) and into the *myc* gene in breast carcinoma (Morse *et al.*, 1988) have been implicated in human carcinogenesis.

1.2 Spermatogenesis

The development of germ cells in male organisms is a good modeling system for studying gene expression and epigenetic regulation. This is due to the massive chromatin remodeling and gene regulation that occurs during the spermatogenic process. Spermatogenesis is the highly regulated developmental process where spermatogonia undergo meiosis and develop into mature spermatozoa (Hess and Renato de, 2008). Spermatogenesis takes place in the seminiferous tubules of the testes, where the germ cells are sequencially organized into several layers containing the different cell types found in the spermatogenic process (Figure 2). Successful formation of mature spermatozoa depends on a succession of signals and cues provided by the local environment (Mruk and Cheng, 2004; Verhoeven, 1992).

1.2.1 The spermatogenic process

The formation of spermatozoa in the testis involves three phases; mitotic proliferation of spermatogonia, meiosis in which the diploid spermatocytes undergo two divisions to form four haploid spermatids, and spermiogenesis which involves a stepwise maturation of the

spermatids to mature spermatozoa. After the meiotic proliferation of spermatogonia, the developing germ cells enter meiosis and become spermatocytes (Figure 2). During prophase I, the spermatocytes go through four substages: leptonema, zygonema, pachynema and diplonema in a highly organized, sequential manner that involves homologous chromosome pairing, synaptonemal complex formation and meiotic recombination (Hess and Renato de, 2008). These processes are followed by two subsequent cell divisions, which reduce the chromosome number by half and produce four haploid spermatids.

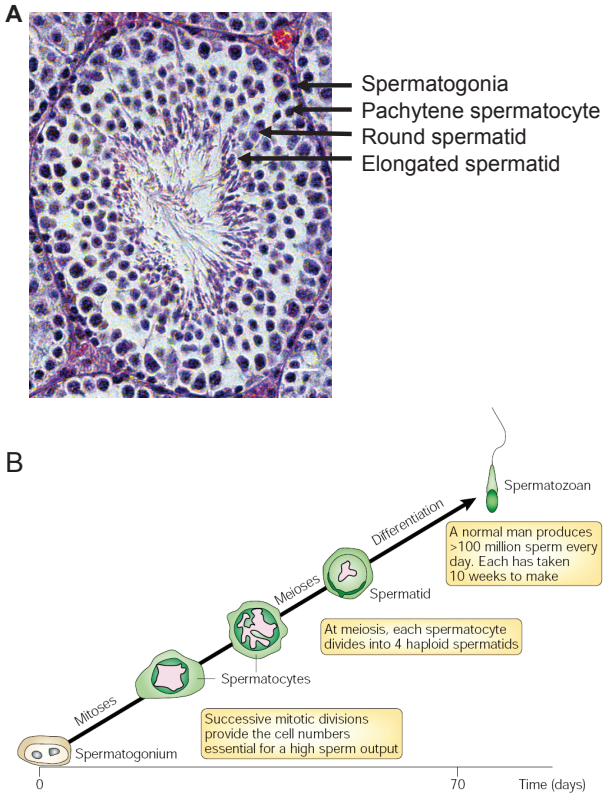


Figure 2. Testicular arrangements and the production of male germ cells. A) HE stained cross section of a stage VII-VIII tubuli from mouse testis containing spermatogonia, pachytene spermatocytes, and round and elongated spermatids (Picture by Kari Furu). B) Schematic diagram of spermatogenesis. During the spermatogenic process, the mammalian germ cell undergoes mitosis, meiosis and structural remodelling into the mature spermatozoa. The process takes 75 days in humans, 35 days in mice (Adapted with permission from Cooke *et al.*, 2002).

After meioses, the spermatids start spermiogenesis, during which they undergo acrosome development and flagellar formation during the round spermatid stage, followed by nuclear condensation and polarization in the elongated spermatid stage (Hess and Renato de, 2008; Carreau and Hess, 2010). At this stage, the histones in the chromatin are replaced with

protamines (Figure 4) resulting in tightly packed chromatin. As a result of the DNA compaction, sperm DNA occupies only 5% of the space occupied by DNA in somatic cells (Monesi, 1964; Ward and Coffey, 1991). This DNA is transcriptionally inactive and proteins will now be produced from mRNAs that have been transcribed and stored during earlier stages of spermatogenesis (Braun, 2000; Monesi, 1964).

In mice, there are 12 stages in the production of male gamets (Figure 3) that together constitutes a cycle generating mature spermatozoa. Each step in the cycle is characterized by a well defined set of cells that synchronously proceed through the spermatogenic process. The meiotic divisions occur in stage XII in mouse and are completed in approximately one day (Hess and Renato de, 2008). The entire process from mitotic division of the spermatogonia to spermatozoa release takes about 35 days in mice and 75 days in humans (Cooke and Saunders, 2002).

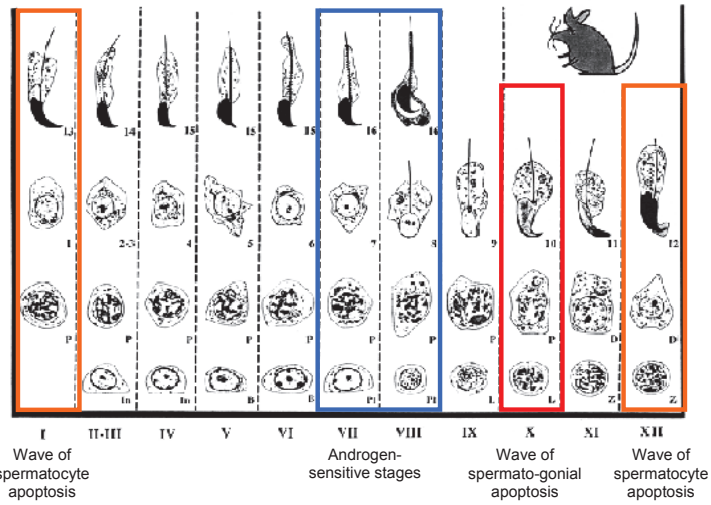


Figure 3. Diagram of the 12 stages for the production of spermatozoa in mice. Each of the 12 stages of spermatogenesis occupies a small segment in the seminiferous epithelium and the maturation continues in each stage. The complete series of stages constitutes a cycle, resulting in the production of mature spermatozoa. The androgen-sensitive stages (blue) and the stages normally containing apoptotic cells (red and orange), are indicated. (Adapted from Russell *et al.*, 1990)

1.2.2 The Androgen Receptor

Androgens and the androgen receptor (AR) are central to the control of spermatogenesis. In testis, the AR is expressed in Leydig cells, the peritubular myoid cells and in the Sertoli cells (SCs) (Kimura *et al.*, 1993; Bremner *et al.*, 1994). The SCs are somatic cells that nurture and support the developing germ cells. They communicate with adjacent cells through specialized cell junctions called gap junctions, ectoplasmic specializations (ES), and tubulobulbar

complexes, through which nutrients and locally produced signaling molecules can be transported (Verhoeven, 1992; Mruk and Cheng, 2004; Tan *et al.*, 2005). The AR functions as a ligand-dependent transcription factor and regulates the expression of an array of androgen-responsive genes, including the genes encoding PSA, c-myc, probasin, KGF, p21 and maspin (Eder *et al.*, 2003; Jiang and Wang, 2003). Binding of AR to androgens results in the formation of homodimer complexes that bind to androgen-response elements (AREs) in the promoter regions of target genes.

The male *AR* knockout mice exhibit a typical female external appearance (Tan *et al.*, 2005). The impact of lacking AR in Sertoli cells mainly affects the ability of the SC to support and nurture germ cells, leading to spermatogenic arrest at the diplotene stage prior to the first meiotic division (Tan *et al.*, 2005). The impact of lacking AR in Leydig cells mainly affects steroidogenic functions leading to arrest of spermatogenesis at the round spermatid stage (Tsai *et al.*, 2006). The deletion of AR gene in mouse germ cells, however, does not affect spermatogenesis and male fertility (Johnston *et al.*, 2001).

At least four steps in germ cell development are controlled by androgen action: spermatogonial differentiation, progression through meiosis, spermatid adhesion and development, and spermiation (Verhoeven, 1992). Immunohistochemical studies show cyclic changes in AR and androgen concentration in Sertoli cells, with peak expression at stage VII and VIII, respectively (Bremner *et al.*, 1994). Stage VII-VIII spermatocytes and round spermatids depend upon androgens to form (Russell and Clermont, 1977), and these stages are therefore called the androgen-sensitive stages (Figure 3). The detachment of the spermatids into the cauda epididymus through spermiation, also takes place here, and probably occurs through an AR-dependent breakdown of the adhesion junctions that have attached the spermatids to SC (Verhoeven *et al.*, 2010),.

1.3 piRNAs

Epigenetically stable repression of transposable elements requires CpG methylation in mammals (Bourc'his and Bestor, 2004). The DNA in primordial germcells (PGCs), however, undergo a wave of DNA demethylation right after their emergence during embryogenesis, rendering the genome vulnerable to transposon-induced mutations. Silencing of transposable elements during this period is, at least in part, governed by the PIWI protein family and a group of small non-coding PIWI-interacting RNAs (piRNAs) (Aravin *et al.*, 2008; Aravin *et al.*, 2007a). In *Drosophila*, 75% of piRNAs are derived from repetitive sequences, and mutations in each of the three members of the PIWI family (Piwi, Aubergine (AUB) and AGO3) lead to transposon activation in germ cells, indicating that they have a non-redundant role in TE silencing (Kalmykova *et al.*, 2005; Vagin *et al.*, 2004; Aravin *et al.*, 2001; Li *et al.*,

2009). In mice, several studies have shown that the PIWI-piRNA pathway is involved in transposon silencing during the wave of TE derepression in testes between 15.5 dpc to 3 dpp (Aravin *et al.*, 2008; Aravin *et al.*, 2007b; Carmell *et al.*, 2007). In addition to this role in TE repression, PIWI-piRNA complexes are also thought to be involved in transcriptional and post-transcriptional mRNA control (Grivna *et al.*, 2006a).

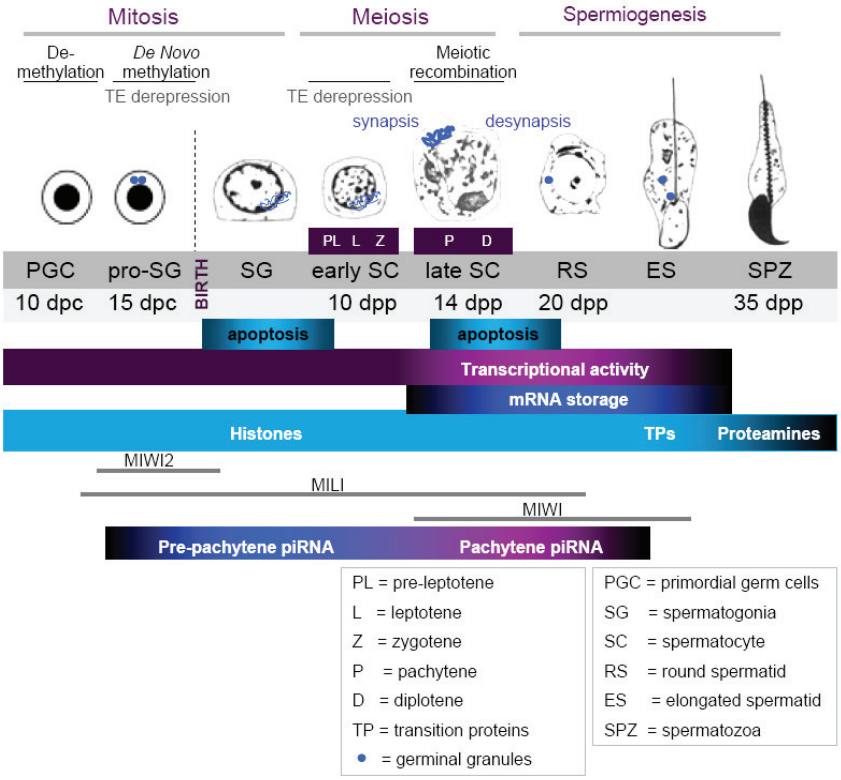


Figure 4. Schematic presentation of the main processes involved in spermatogenesis and the PIWI-piRNA pathway in mice. The steps of germ cell differentiation and temporal expression of piRNAs and related proteins are indicated. The waves of methylation/demethylation and TE derepression, as well as the stage-specific apoptosis of spermatogonia and spermatocytes are also shown (see text for details). (Germ cell images adopted from Russell *et al.*, 1990)

1.3.1 PIWI proteins

piRNAs comprise a large and complex class of small non-coding RNAs. They are defined by size (24-31 nt) and their association with PIWI proteins (Grivna *et al.*, 2006a). The PIWI proteins were first described in *Drosophila melanogaster* and are a subclade of the Argonaute (AGO) proteins, a group of well-described proteins involved in small RNA-mediated silencing complexes (Cox *et al.*, 1998). *Drosophila* Piwi is required for the self-renewal of

germline stem cells. In *Piwi* mutant flies, these cells are depleted and the gonads contain only a few egg chambers or sperm bundles, resulting in infertility in both female and male animals (Cox *et al.*, 1998; Lin and Spradling, 1997). In mice, three PIWI proteins have been identified; Miwi, Mili and Miwi2. Unlike other small non-coding RNAs like miRNAs, both piRNAs and the Miwi-proteins seems to be testis specific (Girard *et al.*, 2006; Aravin *et al.*, 2006). Deletion of any of the three murine PIWI homologs render the male mice infertile and lead to loss of DNA methylation in TEs (Aravin *et al.*, 2008; Kuramochi-Miyagawa *et al.*, 2008). Male *Mili*- and *Miwi2*-knockout mice display meiotic arrest in the developing germ cells (Carmell *et al.*, 2007; Kuramochi-Miyagawa *et al.*, 2004), whereas germ cells in *Miwi*-null testis undergo apoptosis before the onset of spermatid elongation in stage IX (Deng and Lin, 2002).

Each of the three mouse PIWI proteins bind a specific subset of piRNAs, have different expression patterns, and play different roles in the piRNA pathway (Figure 4; Table 1). The Mili protein is expressed in the PGCs and can also be detected in adult testes until the round spermatid stage (Figure 4 and 5), whereas Miwi2 can be detected from 15.5 dpc to 3 dpp (Aravin *et al.*, 2008). The narrow time-window of Miwi2 expression corresponds to a wave of *de novo* methylation (Carmell *et al.*, 2007), and Mili and Miwi2 are involved in transposon repression and methylation of transposable elements during this period (Aravin *et al.*, 2007b; Aravin *et al.*, 2008; Carmell *et al.*, 2007). Miwi is expressed during pachynema and throughout the haploid round spermatid phase in meiosis (Figure 4 and 5) (Grivna *et al.*, 2006b). Unlike the now relatively well-described relation between Mili/Miwi2 and transposon repression, relatively little is known about the function of Miwi and its associating piRNAs.

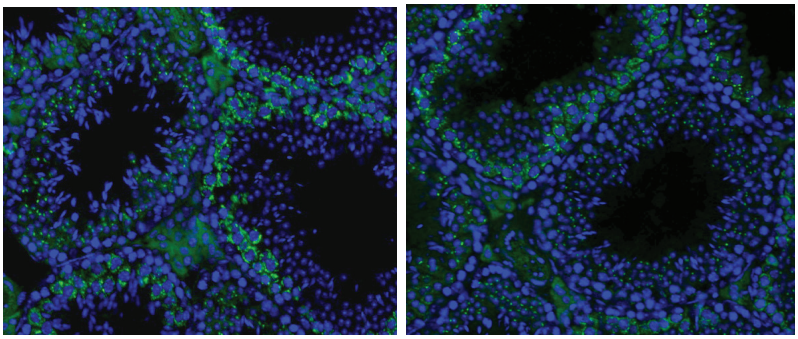


Figure 5. Mili and Miwi distribution patterns in testis. Adult mouse testis stained with anti-Mili (left) and anti-Miwi (right), both in green. Mili is expressed in the cytoplasm of spermatogonia, spermatocytes and round spermatids, whereas Miwi is expressed in spermatocytes from the pachytene stage and throughout the round spermatid stage. Both particularly localizes to the chromatoid body. The sections are counter-stained with DAPI (blue). (Pictures by Kari Furu, unpublished data)

1.3.2 piRNA subgroups and roles in transposon regulation in mice

There are two distinct subsets of mammalian piRNAs that share molecular characteristics, but differ in length, genomic origin, PIWI binding partners, expression pattern, biosynthesis, and function. All piRNAs are between 24-31 nt long, harbor a 5' monophosphate and 3' hydroxyl group, and share a preference for uracile in the first base at the 5' end (5'U) (Beyret *et al.*, 2012; Gan *et al.*, 2011). Mili and Miwi2-associated piRNAs are called pre-pachytene or secondary piRNAs (Table 1). These piRNAs are typically 24-28 nt long, and often have an adenine in the 5' 10 position (5'10A). In prenatal testes, pre-pachytene piRNAs are highly abundant, and about 50% are enriched in transposon sequences (Aravin *et al.*, 2008) (Gan *et al.*, 2011; Ma *et al.*, 2009). Here, DNA methylation and TE-targeting Mili/Miwi2-piRNA-complexes stably silence transposons, reducing the risk of genome instability. Cytosolic Mili uses its endonuclease activity to create piRNAs that inactivate transposon mRNAs (Aravin *et al.*, 2007b; De Fazio S. *et al.*, 2011), whereas nuclear Miwi2 initiates transcriptional silencing by promoting *de novo* DNA methylation of transposon promoter regions (Kuramochi-Miyagawa *et al.*, 2008) (Figure 6).

The second subset of piRNAs, the pachytene or primary piRNAs, are associated with Mili and Miwi, and are 29-34 nt long (Table 1). These piRNAs can only be found in adult animals, and whereas prenatal mouse testes contain a relative abundance of piRNAs mapping to transposonal sequences, piRNAs in adult animals are mostly derived from non-transposonic regions (Aravin *et al.*, 2006; Aravin *et al.*, 2007b; Gan *et al.*, 2011). The vast majority of pachytene piRNAs in mice can be mapped to intergenic sequences (Gan *et al.*, 2011), indicating that the majority of piRNAs in adult testes may function independently of transposon regulation. It is thought that pachytene piRNAs are processed from long, single-stranded RNA precursors that are transcribed from particular DNA sequences known as piRNA clusters (Figure 6; Table 1) (Aravin *et al.*, 2006; Brennecke *et al.*, 2007; Girard *et al.*, 2006; Ragan *et al.*, 2012). Primary piRNA synthesis most likely includes formation of the 5' and 3' ends by unknown endo- and exonucleases (Kawaoka *et al.*, 2011; Pillai and Chuma, 2012), the biogenesis of these molecules are still largely unknown. The role of the PIWI proteins in primary piRNA biogenesis is also unknown, although recent findings indicate that both Mili and Miwi bind the precursor transcripts (Vourekas *et al.*, 2012). The biogenesis of piRNAs is thought to be independent of the nuclease activity of the PIWI proteins, but that they rather determine the length of the piRNA molecules (Siomi *et al.*, 2011; Vourekas *et al.*, 2012). piRNA sequences are not conserved between species, but surprisingly, there is a significant conservation of the genomic locations of the piRNA clusters (Betel *et al.*, 2007).

Secondary piRNAs are probably synthesized through a mechanism called the ping-pong cycle (Brennecke *et al.*, 2007). Here, primary piRNAs synthesized from a transposon-rich cluster

recognise their complementary target, such as a transposon mRNA (Figure 6; Table 1), resulting in cleavage of the transcript 10 nt from the 5' end of the piRNA. This cleavage inactivates the mRNA and creates new piRNAs. These secondary piRNAs will, in turn, target complementary sequences, such as the transcript from the original cluster, creating more primary piRNAs. This is thought to occur in the cytoplasm in prenatal testes and to involve Mili and Miwi2 (Aravin *et al.*, 2008). There is, however, no evidence that the piRNAs present in adult testes are synthesized by this mechanism (Beyret *et al.*, 2012).

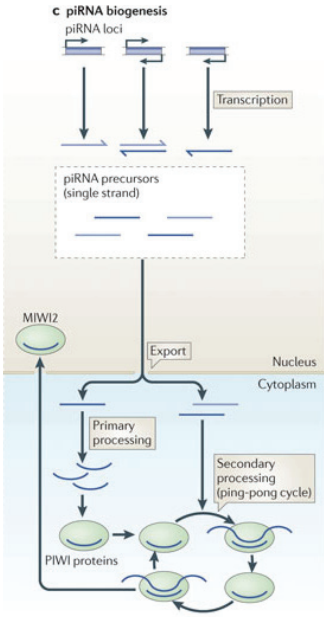


Figure 6. piRNA processing mechanisms in mice. piRNAs are processed from single-stranded RNA precursors that are transcribed largely from particular intergenic regions called piRNA clusters. Primary piRNAs are produced through processing of long primary piRNA transcripts. This occurs in the cytoplasm and the resulting piRNAs (24–31 nt in length) are loaded onto mouse PIWI proteins. The 3′ end of piRNAs is formed by an unknown nuclease followed by 2′-O-methylation. In mouse prospermatogonia, primary piRNAs are processed from transposable elements and can be amplified through the ping-pong pathway. These piRNAs bind Mili in the cytoplasm and the Mili-piRNA complexes bind and cleave anti-sense transcripts, generating 5′ ends of secondary piRNAs that are loaded onto Miwi2. The Miwi2-piRNA complexes, in turn, cleave sense transcripts, generating more primary piRNAs. (Adapted with permission from Siomi *et al.*, 2011)

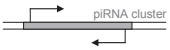
Two studies published in 2011 show that Mili and Miwi have endonuclease activity that is guided by their piRNA partner, leading to cleavage and inactivation of transposon transcripts even in adult male mice (Reuter *et al.*, 2011; De Fazio S. *et al.*, 2011). A need for transposon repression through this pathway in adult animals is likely, as active transcription of some transposons in adult testes is well known (Branciforte and Martin, 1994). Consistent with this observation, a transient relaxation of transposon silencing at the onset of meiosis is observed (Soper *et al.*, 2008). This increased transposon activity drops from mid-pachynema,

coinciding with the appearance of pachytene piRNAs and the intermitochondrial cement. This indicates a role for the PIWI-piRNA pathway in TE repression also in adult testes.

1.3.3 The germinal granules and their role in the PIWI-piRNA pathway

Male germ cells are characterized by distinct RNA- and protein-rich cytoplasmic structures called germinal granules, or nuage (Figure 4) (Eddy, 1975). Germinal granules are evolutionary conserved ribonucleoprotein (RNP) complexes, i.e. complexes composed of one or more proteins and a small, stable RNA molecule. Different forms of germinal granules have been identified (Chuma *et al.*, 2009), of which the most prominent are the intermitochondrial cement (IMC) found in spermatogonia (including in the pro-spermatogonia of the mouse fetus) and all germ cell stages up to the second meiotic division (Chuma *et al.*, 2009), and the chromatoid body (CB) in round and elongated spermatids.

Table 1. Overview of piRNA classes and function

Para-meters \ Type of piRNA	Pachytene piRNAs Type1/ Primary piRNAs	Pre-pachytene piRNAs Type2/ Secondary piRNAs
Length	29-31 nt	24-28 nt
Sequence	5'U + 10N	5'U + 10A
Binding partner	MIWI + MILI	MIWI2 + MILI
Cell type	Pachytene + round spermatids	Spermatogonia - early pachytene
Cell localization	Cytoplasm, particularly to **CB** (both Miwi and Mili are CB components)	**Cytoplasm and nucleus**, enriched in the IMC (both Mili and Miwi2 localize to ICM)
Age of primary expression	Adult testis	Embryos and newborns
Assumed function	Transcription/Translation control???	Transposon repression
Synthesis	Primary processing of piRNA precursor	mRNA from transcribed transposon and ping-pong cycle
DNA origin	Single strand transcript	Double strand transcript
Orientation	Sense and anti-sense	Sense (sense and anti-sense created through ping-pong cycle)
Mapping	 **Intergenic-associated** (80% in adults, 20% in 10 dpp) (from piRNA clusters) mRNA-associated (5% in adults, 35% in 10 dpp) (creates sense piRNAs) Intronic-derived (10% in adults, 15% in 10 dpp) CB = Chromatoid Body ICM = intramitochondrial cement	 **Repeat-associated** (5% in adults, 30% in 10 dpp) **MIWI2** is expressed from 15.5 dpc-3 dpp (gonocytes) **MILI** is expressed from 12.5 dpc in embryos (PGC) and in germ cells up to round spermatid stage **MIWI** is expressed in pachytene and in RS+ES (from 14 dpp)

Work performed in lower organisms such as *Drosophila*, indicate that the nuage plays an important role in the piRNA machinery (Meikar *et al.*, 2011). In these organisms, the germinal granules constitute ribonucleoprotein (RNP) particles containing maternal mRNA required for germ cell specification, indicating that they are important for determining the fate of cellular transcripts (Leatherman and Jongens, 2003). In higher organisms, these structures are likely to have a specialized role in RNA regulation, probably linked to small RNA-mediated gene silencing.

Mili and Miwi2 both localize to the IMC (Aravin *et al.*, 2009), and in pro-spermatogonia, Mili, Miwi2 and other nuage-associated proteins assist in the piRNA pathway to mediate the generation and amplification of piRNAs that are involved in transposon repression (Aravin *et al.*, 2009). A number of proteins, like Gasz, are associated with ICM. Gasz co-localizes with Mili in the ICM and is essential for correct expression of Mili and the production of repeat-associated piRNAs (Ma *et al.*, 2009).

The CB contains both Mili and Miwi, and is loaded with RNA and piRNAs (Meikar *et al.*, 2010). It is therefore possible that this structure is involved in mRNA regulation (Parvinen, 2005), possibly through a mechanism involving piRNAs. Other proteins, like Maelstrom, also localize to the CB. Maelstrom interacts with both Mili, Miwi and unsynapsed chromosomes in spermatocytes, indicating that Maelstrom may function in a piRNA-associated pathway for meiotic silencing of unsynapsed chromatin (MSUC) (Costa *et al.*, 2006).

1.3.4 Possible piRNA functions in translational control and DNA repair

piRNAs are highly abundant in adult testes, indicating an important biological function. However, despite intense research over the past few years, the precise function of pachytene piRNAs is still largely unknown. As pachytene piRNAs originate from clusters spread throughout the genome, one of the functions may be patronage of their respective loci, as seen for scnRNAs in *Tetrahymena* (Mochizuki and Gorovsky, 2004), but other functions such as post-transcriptional regulation and roles in epigenetic programming cannot be ruled out.

Miwi and piRNAs associates with polysomes, indicating a potential role for Miwi and piRNAs in translational control (Grivna *et al.*, 2006b; Kotaja *et al.*, 2006; Meikar *et al.*, 2010; Unhavaithaya *et al.*, 2009). Here, Miwi associates with mRNAs and eIF4E, an mRNA cap binding protein with central roles in translational control (Grivna *et al.*, 2006b). *Miwi* mutant mice have decreased levels of mRNAs that are targets of the spermatogenic transcriptional activator CREM (cAMP-responsive element upmodulator) (Deng and Lin, 2002). Recent findings indicate that Miwi binds mRNAs essential for spermiogenesis without piRNA guides, thus participating in the translational repression occurring in postmeiotic cells (Vourekas *et al.*, 2012). Similarly, studies indicate that Mili may positively regulate

translation and that such regulation is required for germline stem cell self-renewal (Unhavaithaya *et al.*, 2009). In *Mili* mutant testes, the level of protein expression is decreased by 60% without a significant change in mRNA levels. In addition, PIWI proteins interact with Tudor domain–containing proteins, such as Tdrd1 and Tdrd9 (Chuma *et al.*, 2003; Reuter *et al.*, 2009; Shoji *et al.*, 2009). Disruption of either protein leads to male-specific sterility showing meiosis defects and retrotransposons upregulation similar to what has been observed in *Mili* and *Miwi2* null mice (Shoji *et al.*, 2009; Reuter *et al.*, 2009). These results indicate a role for the mouse PIWI homologs in translational control, and that they form effector complexes that regulate mRNA translation and/or mRNA stability.

Pachytene piRNAs associate with Mili and Miwi in the cytoplasm, particularly in the ICM and CB, but they have also been shown to localize to the nucleus (Beyret and Lin, 2011; Marcon *et al.*, 2008). Here, piRNAs localize to a structure called the dense body, which is found in pachytene and diplotene spermatocytes and is associated with the sex chromatin, indicating potential roles in epigenetic regulation and chromatin remodeling. In *Drosophila*, Piwi protein associates with chromatin and directly interacts with Heterochromatin Protein 1 (HP1) (Pal-Bhadra *et al.*, 2004), indicating a role for the PIWI-piRNA pathway in heterochromatin assembly.

Mutations in any of the three *Piwi* genes reduces DNA methylation in mouse testes (Aravin *et al.*, 2008; Kuramochi-Miyagawa *et al.*, 2008). In both *Drosophila* and mice, lack of PIWI proteins leads to increased DNA damage as measured by γ-H2Av foci (a H2A.X variant present at sites of dsDNA damage) (Klattenhoff *et al.*, 2007), indicating a possible role for PIWI proteins in dsDNA break repair (Kuramochi-Miyagawa *et al.* 2004, Carmell *et al.* 2007). Consistent with this hypothesis is the presence of RecQ1 in RIWI complexes in rat (Lau *et al.*, 2006). These findings indicate that PIWI-piRNA complexes may also have a role in maintaining DNA integrity. It is possible that activation of transposons is involved in this increased rate of DNA damage, but the nature of the relationship between DNA damage and transposon mobility is currently unclear.

1.4 The AlkB homologs

The AlkB family is a group of non-heme Fe(II)/α-ketoglutarate (α-KG)-dependent dioxygenases that catalyzes a wide range of biological oxidations (Falnes *et al.*, 2002; Ringvoll *et al.*, 2006; van den Born *et al.*, 2011). The *Eschericia coli* (*E. coli*) repair enzyme AlkB demethylates alkylated DNA lesions such as 1-methyladenine (1meA) to adenine and 3-methylcytosine (3meC) to cytosine in a reaction generating succinate and formaldehyde through a two-step mechanism (Figure 7) (Falnes *et al.*, 2002; Kataoka and Sekiguchi, 1985; Trewick *et al.*, 2002). First, AlkB oxidizes the methyl group, creating an unstable

hydroxymethyl group in a reaction that uses α-ketoglutarate as a substrate. α-KG is reduced to succinate, whereas Fe^{2+} is required as a co-factor. Second, the hydroxymethyl group is spontaneously released as formaldehyde, restoring the original base.

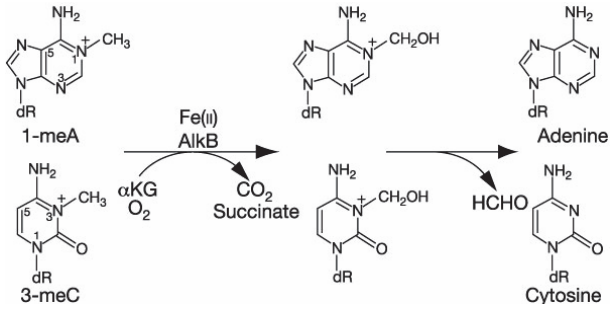


Figure 7. The AlkB mechanism. AlkB removes cytotoxic methyl groups from DNA through a mechanism involving oxidative demethylation using α-ketoglutarate as a substrate and Fe2+ as co-factor. The oxygen in the spontaneously released formaldehyde is derived from O2 (With permission from Trewick *et al.*, 2002).

AlkB homologs are found in many organisms, including *C. elegans*, *S. pombe*, *D. melanogaster* and in many viruses (Aravind and Koonin, 2001). In mammals, nine AlkB homologs have been identified; ALKBH1-8 (Kurowski *et al.*, 2003) and the somewhat less-related obesity-associated FTO (fat mass- and obesity-associated) protein (Gerken *et al.*, 2007). All homologs contain the conserved iron- and 2-oxoglutarate dioxygenase domain, but the nucleic acid binding domain does not share sequence similarity, indicating that some of the ALKB homologs may target molecules other than nucleic acids. AlkB members with known functions are involved in a broad spectrum of cellular processes including maintenance of genome integrity, epigenetic regulation and modulation of translation through tRNA modifications.

ALKBH2 and ALKBH3 are similar to *E. coli* AlkB in that they efficiently reverse damaged bases in nucleic acids in the presence of iron and 2-oxoglutarate *in vitro* by demethylation (Ringvoll *et al.*, 2008; Ringvoll *et al.*, 2006; Duncan *et al.*, 2002; Falnes, 2004). ALKBH8 contains an additional C-terminal SAM-dependent tRNA methyltransferase domain, required for the final step in the biogenesis of 5-methoxycarbonylmethyluridine (mcm^5U) (Songe-Moller *et al.*, 2010; Fu *et al.*, 2010). mcm^5U is involved in decoding the UGA stop-codon for the synthesis of selenoproteins. This indicates a role in translational decoding. ALKBH8 also contains an N-terminal DNA/RNA recognition motif in addition to the AlkB domain. The AlkB domain catalyzes the hydroxylation of mcm^5U to $mchm^5U$ at the wobble position of tRNA (Fu *et al.*, 2010; van den Born *et al.*, 2011). The ninth AlkB homolog, the obesity-associated FTO protein, has been shown to demethylate 3-methylthymine (3meT) and 3-methyluracil (3meU) in single stranded DNA and RNA *in vitro* (Gerken *et al.*, 2007; Jia *et*

al., 2008). Recently, it was shown that FTO is an RNA demethylase that oxidatively removes the 6meA modification in mRNA *in vitro* and *in vivo* (Jia *et al.*, 2011).

The function of ALKBH4, 5, 6 and 7 are currently unknown. It is known that *E. coli* AlkB and ALKBH2, 3 and 8, have high pI values and positively charged surfaces, and that they have nucleic acids as substrates (Sedgwick *et al.*, 2007). Based on the similarly high pI values of ALKBH5 and 6, it has been suggested that these proteins also work on nucleotides or nucleic acids. The relatively low pI of ALKBH1, 4 and 7 indicate that these proteins may have other main targets than nucleic acids. Jumonji (JmjC)-domain containing histone demethylases all have low pI values and remove methyl groups from histones using the same mechanism as *E. coli* AlkB. It has therefore been suggested that Alkbh1, 4 and 7 are involved in protein/histone demethylation (Loenarz and Schofield, 2008; Sedgwick *et al.*, 2007).

1.4.1 Alkbh1

The human *ALKBH1* gene was the first AlkB homolog to be discovered in mammals (Wei *et al.*, 1996). The ALKBH1 protein shows a high degree of homology with the *E. coli* AlkB, but the precise function remains unknown. Several studies have been published reporting diverse enzymatic functions and subcellular localization of the ALKBH1 protein (Muller *et al.*, 2010; Tsujikawa *et al.*, 2007; Westbye *et al.*, 2008). Recombinant truncated Alkbh1 enzyme may demethylate 3meC in single stranded DNA and RNA *in vitro* (Westbye *et al.*, 2008). Another *in vitro* study found that ALKBH1 possesses a DNA lyase activity, in which AP-specific DNA cleavage occurs in a Fe(II)- and 2OG independent manner (Muller *et al.*, 2010). A lyase activity has also been described in the ALKBH1 homolog Abh1 in *Schizosaccharomyces pombe* (Korvald *et al.*, 2012), but it remains unclear whether these activities are physiologically relevant. In 2008 a paper on Alkbh1 was published by Pan et al, where a gene-targeting study in mice showed that Alkbh1 localizes to nuclear euchromatin (Pan *et al.*, 2008). Their study demonstrated impaired placental trophoblast lineage differentiation in *Alkbh1*$^{-/-}$ mice, and a strong interaction of Alkbh1 with Mrj, an essential placental gene that mediates gene repression by recruitment of class II histone deacetylases (HDAC). This study thus indicates a function for Alkbh1 in epigenetic regulation during development.

1.5 BTB/POZ zinc finger proteins

The Broad complex, Tramtrack, Bric a Brac or Pox virus and Zinc finger (BTB/POZ) zinc finger family is a group of transcription factors found in many organisms from viruses throughout eukaryotes. These proteins carry a BTB/POZ domain in their N-terminal region and one or more zinc fingers in the C-terminus. The name BTB in the BTB/POZ domain is named after proteins in *Drosophila* containing this domain. Broad complex is an early response element in ecdysone-induced metamorphosis (DiBello *et al.*, 1991), Tramtrack is

crucial to embryogenesis and cell fate determination in the eye (Xiong and Montell, 1993), and Bric a brac is required for pattern formation along the proximaldistal axis of the leg and antenna (Godt *et al.*, 1993). In vertebrates, proteins containing the BTB/POZ domain contribute to transcription repression (Melnick *et al.*, 2000; Muto *et al.*, 1998), cytoskeleton regulation (Ziegelbauer *et al.*, 2001), protein ubiquitination/ degradation (Xu *et al.*, 2003), and even oncogenesis (Shaffer *et al.*, 2000). Studies have elucidated that the BTB/POZ domain mediates homomeric or heteromeric dimerisation (Bardwell and Treisman, 1994), and promotes transcriptional repression through the recruitment of co-repressor proteins such as the nuclear co-receptor/silencing mediator for retinoid and thyroid hormone receptors (N-Co-R/SMRT) complex (Dhordain *et al.*, 1997). A histone deacetylase can subsequently be recruited by these co-repressors to the promoters of target genes, leading to gene silencing (Chapter 1.1.2).

In addition to the BTB/POZ domain, the BTB/POZ zinc finger (BTB/POZ-ZF) proteins contain one or several zinc fingers in their C-terminal region. Zinc finger motifs are DNA-binding domains defined by a requirement for zinc ions. Several types exists, of which C_2H_2 zinc fingers are the most clearly identified motif with a wide existence in human, animals and plants (Klug and Schwabe, 1995). This motif often occurs in tandem repeats characterized by two cysteine and two histidine residues that coordinate a zinc ion and fold the domain into a finger-like projection, enabling the protein to interact with DNA.

1.5.1 Tzfp

Testis Zinc Finger Protein (TZFP) (also known as ZBTB32, PLZP, ROG and FAZF) is a BTB/POZ-ZF protein thought to play a role in spermatogenesis due to its high expression in testis (Lin *et al.*, 1999; Hoatlin *et al.*, 1999). Mice deficient of Tzfp has previously been generated (Kang *et al.*, 2005; Piazza *et al.*, 2004), revealing that the protein is not embryonic lethal and that mice lacking Tzfp develop without any obvious phenotype. In blood, TZFP is expressed at high levels in the early stages of differentiation but declines during subsequent differentiation into erythroid and myeloid lineages (Dai *et al.*, 2002). Here, disruption of the protein leads to increased T-lymphocyte proliferation, cytokine production and altered hematopoetic stem cell homeostasis (Piazza *et al.*, 2004), as well as altered kinetics of the regulation of genes involved in B-lympocyte maturation into plasma cells (Yoon *et al.*, 2012). Overexpression of the protein is accompanied by accumulation in G1 and increased rates of apoptosis (Dai *et al.*, 2002).

The effects of Tzfp disruption in blood are probably due to its repressive effect on GATA factors (Miaw *et al.*, 2000). GATA is a group of evolutionarily conserved transcriptional regulators that play crucial roles in the development and differentiation of all eukaryotic organisms. TZFP represses GATA-3, resulting in repression of Th cell transactivation and

production of Th cytokines (Miaw *et al.*, 2000). TZFP has also been shown to interact with GATA-2 (Tsuzuki and Enver, 2002). These results suggest a role for Tzfp during the proliferative stages of primitive hematopoietic progenitors, particularly in proliferation and maturation of lymphocytes. Its precise roles in other forms of cell differentiation, however, remain unknown.

TZFP forms a heterodimer with the highly homologous protein Promyelocytic leukemia zinc finger protein (PLZF) (Hoatlin *et al.*, 1999). PLZF plays critical roles in oncogenesis, patterning of the limb and axial skeleton, T lymphocyte development, and maintenance of spermatogonial stem cells (Barna *et al.*, 2000; Costoya *et al.*, 2004; Kovalovsky *et al.*, 2008). TZFP also interacts with Fanconi Anemia Complementation group C (FANCC) (Hoatlin *et al.*, 1999), a member of the Fanconi Anemia (FA) pathway associated with the repair of interstrand cross-links in DNA (Auerbach, 1988; Moldovan and D'Andrea, 2009). Lack of FANCC or other FA members, result in anemia and bone marrow failure. FANCC functions in homologous recombination, which promotes mutational repair of endogenously generated abasic sites (Niedzwiedz *et al.*, 2004).

Tzfp probably exists in at least two isoforms (Miaw *et al.*, 2000; Lin *et al.*, 1999; Ikeda *et al.*, 2007). Isoform 1 consists of 465 amino acids and is predicted to be testis specific, whereas the shorter isoform 2 is expressed in B-and T lymphoid cells and lacks the C-terminal zinc finger motif. This indicates that the longer, testis specific isoform can bind DNA through its zinc finger motif, thereby initiating changes in gene expression. In 2001, Tang and co-workers showed that the zinc finger domain binds to a specific genetic sequence, TGTACAGTGT, located at the upstream flanking sequence of the *Aie1* gene. They called this sequence the Tzfp binding site (tbs) and showed that binding to the gene through this sequence has a repressive effect (Tang *et al.*, 2001), possibly via the recruitment of N-CoR/SMRT complexes.

Based on published studies, it seems clear that TZFP has a regulatory role in blood. The highest expression of this protein is, however, found in testis (Lin *et al.*, 1999; Hoatlin *et al.*, 1999; Miaw *et al.*, 2000), but its role here remains unclear. The observations that TZFP is not restricted to one cell type or organ, the presence of different isoforms, and the possibility that it engages in both protein-protein interaction via its BTB/POZ domain and corepressor recruitment by binding to DNA via its zinc fingers (Hoatlin *et al.*, 1999), raise the possibility that TZFP may play different roles in different cell types.

2 Aims of study

The enzymatic function of *E. coli* AlkB was determined in 2002 when its ability to repair 1meA and 3meC through a direct, oxidative mechanism was described by two individual groups (Falnes *et al.*, 2002; Trewick *et al.*, 2002). Later, it became evident that genes homologues to the *AlkB*-encoding gene are found in many organisms, including in the human and murine genome, all containing the 2OG-Fe(II) oxygenase domain. The roles of the proteins encoded by human *ALKBH1* and murine *Alkbh1* remains elusive, despite their similarities to AlkB in *E. coli*. The studies and results described in this thesis have been performed to elucidate the molecular function of this protein and its potential binding partner TZFP. With the use of animal models and *in vitro* experiments, we have attempted to identify which biological mechanisms these two proteins are involved in.

TZFP is a BTB/POZ-ZF protein thought to have a repressive effect on target genes (Lin *et al.*, 1999; Tang *et al.*, 2001). The protein has been found to be a repressor of GATA in lymphocytes (Miaw *et al.*, 2000), indicating a role in T cell regulation. We had observed that both *Tzfp* and *Alkbh1* are highly expressed in testes, and we therefore set out to identify the role of these two proteins in spermatogenesis.

ALKBH1 had previously been suggested to target protein substrates and, perhaps, to be involved in epigenetic regulation, due to its lack of activity on DNA and RNA substrates (Aas *et al.*, 2003). It was suggested that ALKBH1 may be a histone demethylase, modifying gene activity through altering epigenetic marks at target genes. *Alkbh1* deficient mice enabled us to investigate the phenotype of animals lacking a functional Alkbh1 protein. Murine Tzfp had previously been described as a transcription repressor able to repress the testis specific gene *Aie1* through binding to the Tzfp binding sequence in the gene's promoter (Tang *et al.*, 2001). Therefore, using mutant mice, we sought to identify Tzfp target genes and downstream effects of *Tzfp* deletion in spermatogenesis.

piRNAs, a group of small non-coding RNAs, were first described in mice in 2006 (Aravin *et al.*, 2006; Grivna *et al.*, 2006a). piRNAs are 24-31 nt in length and are specific for the mammalian male germline. They associate with PIWI, a group of spermatogenesis-specific proteins. piRNAs and PIWI proteins have been implicated in transposon control and are linked to transposon methylation in the testes of prenatal and newborn mammals. We investigated a possible role in piRNA regulation for Alkbh1 and Tzfp, and identified a piRNA cluster whose expression is controlled by these two proteins.

3 Summary of Papers

3.1 Paper I: Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects.

In Paper I, we characterize mice lacking *Alkbh*, designed by deleting exon 6 in the *Alkbh1* gene. Mice homozygous for the mutated allele display a non-Mendelian pattern of inheritance, and both homozygous and heterozygous disruption leads to reduction in survival rates. Additionally, the sex-ratio is considerably skewed against female offspring, with one female born for every three to four males. The sex ratio distortion appears to be of paternal origin, and might originate from the pachythene stage of meiosis during spermatogenesis. *Alkbh1* is highly expressed in pachynema, where the gene is upregulated more than 10-fold when compared to spermatogonial expression. In addition, apoptotic spermatids were revealed in 5-10% of the tubules in *Alkbh1*$^{-/-}$ testes.

The deficiency of Alkbh1 causes aberrant expression of Bmp2, 4 and 7 at E11.5 during embryonic development. This is consistent with an incompletely penetrant phenotype of recurrent unilateral eye defects and craniofacial malformations.

Our data suggest that Alkbh1 plays a role in spermatogenesis, and that Alkbh1 is essential for normal sex-ratio distribution and embryonic development in mice. The phenotypic observations indicate that Alkbh1 is involved in differentiation, particularly during spermatogenesis and during embryonic development.

3.2 Paper II: Alkbh1 and Tzfp Repress a Non-Repeat piRNA Cluster in Pachytene Spermatocytes

Here, we discuss the possible role of Alkbh1 and Tzfp in the PIWI-piRNA pathway. piRNAs are small non-coding RNAs specific for the male germline in mammals that interact with a class of proteins called PIWI. PIWI proteins and piRNAs have known functions in transposon silencing in fetal and newborn mice. Both are, however, abundant in adult testes, but their function here remains a mystery.

In this paper, we isolate specific germ cells and perform deep sequencing of small RNA to show that the majority of piRNAs in spermatocytes originates from clusters in non-repeat, intergenic DNA. The regulation of piRNA clusters, and the processing of the primary transcripts into individual piRNAs, is accomplished through mostly unknown processes. We present a possible regulatory mechanism for one such cluster, named cluster 1082B, located on chromosome 7 in the mouse genome. The 1082B precursor transcript and its 788 unique

piRNAs are repressed by the Alkbh1 dioxygenase and the testis-specific transcription repressor Tzfp. We observe a remarkable >1000 fold upregulation of individual piRNAs in pachytene spermatocytes isolated from *Alkbh1* and *Tzfp* deficient murine testes. Repression of cluster 1082B is further supported by the identification of a 10 bp Tzfp recognition sequence contained within the precursor transcript. Downregulation of LINE1 and IAP transcripts in the *Alkbh1* and *Tzfp* deficient mice lead us to propose a potential role for the 1082B-encoded piRNAs in transposon control.

In conclusion; our results suggest that Alkbh1 and Tzfp regulate a primary piRNA transcript encoded by cluster 1082B on chromosome 7 in mice. The primary transcript gives rise to 788 unique piRNAs that has a proposed role in silencing of IAP and L1 transposons.

3.3 Paper III: Tzfp represses the Androgen Receptor in mouse testis

This paper is a description of a novel *Tzfp* deficient mouse and suggests a role for the testis zinc finger protein (Tzfp) in the regulation of several genes and proteins in mouse testes. Tzfp, also known as Repressor of GATA, belongs to the BTB/POZ-ZF family of transcription factors and is thought to play a role in spermatogenesis due to its remarkably high expression in testis. Despite many attempts to find the *in vivo* role of the protein, the molecular function is still largely unknown. Mice lacking Tzfp are viable and fertile, but display an increase in Androgen Receptor (AR) signaling in Sertoli cells and altered expression of several genes in the testis. Our results indicate that Tzfp represses *Gata-1*, *Aie1* and *Fancc*, leading to altered AR signaling, resulting in a reduced number of apoptotic cells in the testicular tubules.

In conclusion; these results indicate that Tzfp is a potent repressor of several genes in mouse testes and that it is part of a complex regulatory system that ensures proper Androgen Receptor signaling during spermatogenesis.

4 Results and Discussion

4.1 Alkbh1

Despite many studies on ALKBH1 function, two of which include *Alkbh1* deficient mice, the molecular role of this protein is still largely unknown. In 1996 it was claimed that ALKBH1 could partially rescue cells from MMS-induced cell death, indicating a role in maintenance of DNA integrity through a mechanism similar to *E. coli* AlkB (Wei *et al.*, 1996). A few years later, a paper reporting that ALKBH1 is able to cleave DNA at abasic (AP) sites through a lyase mechanism, was reported (Muller *et al.*, 2010). A third group reported that ALKBH1 is a mitochondrial protein demethylating 3meC in DNA and RNA (Westbye *et al.*, 2008). The papers all claim that ALKBH1 is an enzyme involved in the repair of nucleic acids, but most of these results have yet to be reproduced by other groups.

4.1.1 *Alkbh1* phenotypes – developmental defects and sex ratio distortions

Our results indicate that Alkbh1 is not a repair protein involved in the reversal of base lesions in nucleic acids, but that it rather is involved in epigenetic regulation during embryonic development and spermatogenesis (Paper I). This challenges previous data that proposes that ALKBH1 is a repair enzyme. Redundancy for ALKBH1 activity in DNA repair pathways is supported by the lack of a cancer and neurodegenerative phenotype in Alkbh1 deficient mouse strains (Paper I) (Pan *et al.*, 2008).

The phenotypic defects of the *Alkbh1*$^{-/-}$ mouse strain indicate a role in differentiation and development of cells in tissues of ectodermal origin (Paper I). Embryos display malformations of the eye, delayed occification (bone formation) of sternum and skull, and neural tube defects like spina bifida and exencephaly. These phenotypic traits have an incomplete penetrance, i.e. they are present in only a subset of *Alkbh1*$^{-/-}$ embryos. This indicates that the phenotypic traits are polygenic, with the trait being modified by the presence or absence of polymorphic alleles at other gene loci (Glazier *et al.*, 2002). The dependence upon modifier genes and gene dosage may lead to irregular patterns of inheritance and might explain the inconsistent phenotypes observed in the *Alkbh1*$^{-/-}$ pups. One candidate group of genes that might act as such modifiers, are the genes encoding the bone morphogenetic proteins (BMPs). The BMPs are a group of proteins within the Transforming Growth Factor beta (TGF-ß) superfamily working upstream of SMAD signaling pathways with important roles in the regulation of bone induction, maintenance and repair (Ripamonti and Reddi, 1992; Ripamonti and Reddi, 1997). During embryogenesis, they regulate dorsal-ventral patterning, establishment of embryonic body plan, apoptosis, differentiation of neural

cells, patterning of limb buds, and epithelial-mesenchymal interactions during organogenesis (Jones *et al.*, 1992; Zou and Niswander, 1996; Paralkar *et al.*, 1992; Shah *et al.*, 1996; Vainio *et al.*, 1993; Cheifetz, 1999). Proper BMP function is dependent upon gene dosage, and upon expression- and function-overlap of the different BMPs (Bandyopadhyay *et al.*, 2006; Ducy and Karsenty, 2000).

We found that several BMPs are aberrantly expressed in *Alkbh1*$^{-/-}$ embryos (Paper I). Interestingly, some of the phenotypes observed in *Alkbh1*$^{-/-}$ mice are similar to those observed in mice models and humans with lack of, or mutations in, genes encoding BMP proteins. The phenotypes typicly exhibit highly variable extraocular features and incomplete penetrance (Wyatt *et al.*, 2010). Eye developmental defects have been observed in mice lacking Bmp4 and Bmp7 (Wordinger and Clark, 2007; Wyatt *et al.*, 2010; Dudley *et al.*, 1995). The eye is derived from the neural tube (neuroectoderm), from which the retina and its associated pigment cell layer develop. Like the *Alkbh1*$^{-/-}$ mice, BMP7 deficient animals also exhibit a number of specific skeletal defects with variable penetrance (Dudley *et al.*, 1995). Bmp2, Bmp4 and Bmp7 are necessary for closure of sutures in the skull (Hall and Miyake, 2000), indicating that the reduced or missing ossification observed in some of the *Alkbh1*$^{-/-}$ embryos could be due to mis-expression of these proteins.

Alkbh1 mutant embryos were found to have reduced viability and a sex-ratio disruption, with 3-4 male pups born for every female. Paternal inheritance of the targeted allele seemed to be more critical than maternal transmission for the survival of offspring, as HetxKO breedings with a KO male resulted in fewer *Alkbh1*$^{-/-}$ pups than breedings with KO females. Such non-Mendelian inheritance can be a result of a segregation distortion of the sex chromosomes during meiosis, as sex chromosomal distorters often lead to biased sex ratios (Taylor and Ingvarsson, 2003). This makes the pairing and silencing of the X and Y chromosomes potential Alkbh1 targets.

Due to the large non-homologous segments in the sex chromosomes, the X and Y chromosomes behave differently from the autosomes during meiosis. This is particularly evident during pachynema when the sex chromosomes form the so-called XY body, which is a diagnostic feature for pachytene spermatocytes (Handel, 2004; Solari, 1974). This structure is characterized by condensed chromatin, meiotic sex chromosome inactivation (MSCI), and restricted chromosome pairing. At the zygotene/pachytene transition there is DSB-independent phosphorylation of H2A.X at the XY chromosomes, triggering chromatin changes that lead to MSCI (Fernandez-Capetillo *et al.*, 2003). In H2A.X deficient cells, autosomal synapsis is normal, but is reduced in the sex chromosomes (Celeste *et al.*, 2002). In addition, these cells lack proper MSCI and the crossing-over frequency is sharply reduced, leading to arrest at early pachynema (Celeste *et al.*, 2002; Fernandez-Capetillo *et al.*, 2003).

Alkbh1 deficient pachytene spermatocytes display normal formation of the XY body, indicating normal segregation of the sex chromosomes (Paper I). This does not exclude the possibility that lack of Alkbh1 results in an epigenetic and silencing defect of the paternal sex chromosomes. More investigation is needed to establish the exact mechanism by which this phenotype is caused.

4.1.2 Other *Alkbh1* deficient mice

As described above, *in vitro* studies have led to the proposal that ALKBH1 could work in nucleic acid repair (Muller *et al.*, 2010; Wei *et al.*, 1996; Westbye *et al.*, 2008). To study the *in vivo* role of the protein, we and one other group have created Alkbh1 deficient mice that challenge these findings, as the phenotypic defects of the mice indicate a role for Alkbh1 in development and cell differentiation (Paper I) (Pan *et al.*, 2008). There are, however, some differences in the phenotypes observed in the two strains regarding the tissues affected, revealing discrepancies between the two mouse models. Pan and co-workers did not report any eye or skeletal abnormalities in their *Alkbh1* deficient mouse, and observed no sex ratio distortions (Pan *et al.*, 2008). However, they found that Alkbh1 interacts with Mrj, a co-chaperone member of the DnaJ protein family widely expressed in the adult and embryonic mouse. Mrj mediates gene repression by recruitment of class II HDACs (Dai *et al.*, 2005). It is most highly expressed in brain and trophoblast cells of the placenta, and is required for normal placental development (Chuang *et al.*, 2002; Hunter *et al.*, 1999). It has also been suggested that Mrj is an essential regulator of neural stem cell self-renewal (Watson *et al.*, 2009), as the *Mrj*$^{-/-}$ mouse displays a phenotype that includes neural tube defects. Interestingly, this phenotype resembles phenotypic traits observed in our *Alkbh1*$^{-/-}$ mouse (Paper I). This may support the finding by Pan *et al.* that Alkbh1 and Mrj are interaction partners working in the same pathways. As Mrj is a co-chaperone protein, it is tempting to speculate that lack of Mrj leads to aberrant Alkbh1 localization, thus preventing Alkbh1 to bind to its target substrates.

The discrepancy between the two models can probably not alone be explained by the fact that the two mutants were generated with different genetic backgrounds (129/SvEv and C57BL/6). More likely, the difference is due to differences in gene targeting strategy. Pan and co-workers deleted Exon 3, resulting in a frame shift and premature stop codon in exon 4. We deleted exon 6 and replaced it with a neomycin cassette, removing the conserved 2OG interaction domain but maintaining the remaining protein domains. Our strategy may have created a truncated variant of Alkbh1 able to bind certain interaction partners. Pan *et al.* however, claim that they have generated a null mouse, with no truncated protein present. Despite the discrepancies in phenotype, the findings suggest that the effect of Alkbh1 deficiency is pleiotropic and dependent on cell type and/or stage of development.

4.1.3 A possible molecular function for ALKBH1 in histone demethylation

The protein family of 2OG oxygenases is a large superfamily involved in a wide range of biological processes (Loenarz and Schofield, 2008). These proteins couple the oxidation of their substrate to the oxidative decarboxylation of 2OG to give succinate and carbon dioxide. In mammals, they oxidize a wide range of substrates and are involved in processes such as fatty acid metabolism, collagen formation and hypoxic signaling (Jenkins and Raines, 2002; Ozer and Bruick, 2007; Coleman *et al.*, 2007; Schofield and McDonough, 2007). In addition, Tsukada and colleagues demonstrated that a 2OG oxygenase from the jumonji domain–containing (JmjC) family is capable of catalyzing a demethylation reaction of tri-, di- and monomethylated lysyl histone peptides (Tsukada *et al.*, 2006). Several JmjC 2OG oxygenases with histones as their substrate have now been identified (approximately 30 in humans).

The findings that JmjC proteins act as histone demethylases, lead to the hypothesis that ALKBH1 may have histones as its substrate rather than DNA. Interestingly, Alkbh1 deficient mice display a phenotype that resembles the pathologies demonstrated in knockout models of histone demethylases. This includes the eye development distortions and craniofacial defects observed in the *Sox3*-null mutants (a HMG box factor) and the phosphatidylserine receptor (*Ptdsr*) deficient mice (a histone arginine demethylase), respectively (Bose *et al.*, 2004; Rizzoti and Lovell-Badge, 2007). This may indicate that ALKBH1 has a role in generating a histone modification pattern necessary to ensure proper gene expression during differentiation and development. Interestingly, Alkbh1 localizes to nuclear euchromatin (Pan *et al.*, 2008) and recent studies suggest that ALKBH1 is a histone H2A dioxygenase (Ougland *et al.*, in press).

4.1.4 A possible role for Alkbh1 in spermatogenesis and male fertility

The *Alkbh1* gene was found to be most highly expressed in testes (Paper I). Additionally, *Alkbh1*$^{-/-}$ animals displayed a decreased testicular weight and an increased rate of apoptosis within the tubules of ageing males when compared to wild-type. This indicates a regulatory role for Alkbh1 in spermatogenesis, possibly in synapsis or apoptosis. Despite this, *Alkbh1*$^{-/-}$ animals were found to be fertile and showed no obvious signs of fertility abnormalities.

Synapsis is the intimate pairing of the parental chromosomes during meiosis (Burgoyne *et al.*, 2009; Roeder, 1997). Vital for correct synapsis formation are the numerous DSB that naturally occur during leptonema in the sister chromatids (Dernburg *et al.*, 1998; McKim and Hayashi-Hagihara, 1998). During zygonema, the axes of each pair of homologous maternal and paternal chromosomes begin synapsis (Figure 4), a process that involves the formation of transverse and axial elements that make up the synaptonemal complex (Burgoyne *et al.*, 2009; Zickler and Kleckner, 1999; Marston and Amon, 2004). When synapsis is completed by the

initiation of the pachytene stage, the DSB are repaired by homologous recombination and some of the breaks will mature into crossovers. After crossing over, the synaptonemal complex breaks down, but the two duplicated homologs remain attached at the points where the crossovers occurred, called the chiasmata (Hirose *et al.*, 2011; Nicklas, 1974).

Errors in the synapsis of homologous chromosomes are associated with a male-biased meiotic impairment (Burgoyne *et al.*, 2009; Turner *et al.*, 2006). This is thought to be a response to persisting DSB in the asynapsed chromosome segments which, in turn, leads to apoptosis during pachynema or during the first meiotic division (de Rooij and de Boer P., 2003). Synapsis and meiotic recombination are thus critical steps in meiosis, and a pachytene checkpoint prevents exit from pachynema when these two processes are incomplete (Roeder, 1997; Shuster and Byers, 1989). Asynapsis during meiotic prophase I also leads to meiotic silencing of unsynapsed chromosomes (MSUC), a phenomenon normally restricted to the sex chromosomes, and results in transcriptional silencing of the affected chromosomal regions (Baarends *et al.*, 2005; Turner *et al.*, 2006; Turner *et al.*, 2005). These phenomena makes it tempting to speculate that Alkbh1 deficiency results in increased asynapsis that, in turn, leads to improper expression of proteins needed to complete meiosis and spermiogenesis, leading to increased apoptosis. As described above, no abnormalities were seen in the staining patterns of γH2A.X, indicating normal formation silencing of the sex chromosomes and absence of abnormal double strand breaks in the autosomes (Paper I). In addition, normal formation of the ynaptonemal complex was observed, as seen by staining pachytene spermatocytes with anti-SYCP1 and -3 (unpublished data), indicating normal synapsis of the autosomes. Thus, synapsis seems not to be affected in the *Alkbh1*$^{-/-}$ animals.

Apoptosis could be the outcome of various pathologies and of a deregulation of spermatogenesis control systems (Gandini *et al.*, 2000). Apoptosis normally takes place in the testes of many animals, and helps to maintain an appropriate number of each germ cell type, as well as to selectively remove damaged cells (Print and Loveland, 2000). Apoptosis is stage-specific, and spermatogonial and spermatocyte death usually occurs at stage X (mainly apoptotic spermatogonia) and stage XII-I (mainly apoptotic spermatocytes) in mice (Figure 3 and 4) (Blanco-Rodriguez *et al.*, 2003; Bartke, 1995). Apoptosis in other stages or cell types is rare in healthy animals, but has been documented in several pathologies. For example, hormonal dysregulation such as lack of testosterone, causes apoptosis in stage VII-VIII spermatocytes (Henriksen *et al.*, 1995), whereas *Dmc1*$^{-/-}$ and *Atm1*$^{-/-}$ spermatocytes undergo apoptosis during stage IV (Barchi *et al.*, 2005).

In the *Alkbh1*$^{-/-}$ testes, apoptosis was observed in spermatocytes and spermatids in 5–10% of the seminiferous tubules of one year old animals. The affected tubules have abnormal cellular organization and contain fewer cells than normal. This reduction in the number of

spermatogenic cells might reflect an indirect effect of prolonged arrest in spermatids in the affected tubules. The pleiotropic and leaky mutant phenotype observed in *Alkbh1*$^{-/-}$ testes, is comparable to several other mouse models with deleted spermatogenesis-related genes. Disruption of the regulatory genes *Dazla* and *Mvh*, (Ruggiu *et al.*, 1997; Tanaka *et al.*, 2000), the cell cycle regulator *cyclin A1* (Liu *et al.*, 1998) and the transcription factor *A-myb* (Toscani *et al.*, 1997) all affect spermatogenesis at multiple stages, often with a variable range of defects and incomplete penetrance. This indicates that Alkbh1 could regulate proteins involved in the differentiation and proliferation of spermatogenic cells.

As BMPs seem to be involved in the phenotypic defects in *Alkbh1*$^{-/-}$ embryos, it is possible that this group of proteins is also involved in the testicular defects observed in these animals. BMPs influence germ cells and somatic cells throughout fetal and postnatal life, and Bmp4, Bmp8a and Bmp8b are required for the maintenance of adult spermatogenesis *in vivo* (Hu *et al.*, 2004; Zhao and Hogan, 1996). *Bmp8a* and *Bmp8b* are both expressed in the testes, specifically in stage 6-8 round spermatids, suggesting that they regulate aspects of cell proliferation, survival and/or differentiation during spermatogenesis (Zhao and Hogan, 1996). *Bmp8b* mutant males exhibit reduced germ cell proliferation in juvenile animals, and a significant increase in apoptotic spermatocytes in adult animals (Zhao *et al.*, 1996). The defects render the animals infertile, probably due to an effect of Bmp8b in male germ-cell proliferation during early puberty, and in germ-cell survival and fertility in adult animals. *Bmp4* heterozygous males also show compromised fertility due to degeneration of germ cells, reduced sperm counts, and decreased sperm motility (Hu *et al.*, 2004). The *Bmp* mutant mice thus display similarities with *Alkbh1*$^{-/-}$ mice, indicating that the defects in *Alkbh1*$^{-/-}$ testes could be due to dysregulation of BMPs in male germ cells. An increase in spermatogenic cell death can be a result of several defective mechanisms, so a possible link between the phenotype in *Alkbh1*$^{-/-}$ males and a testicular dysregulation of Bmp's must be validated through further studies.

In addition to a possible role for Alkbh1 in testes through regulation of BMP expression, the phenotype observed in *Alkbh1*$^{-/-}$ testes might be related to the putative role of Alkbh1 as a histone demethylase (Ougland *et al.*, in press). Yet, testis sections from *Alkbh1*$^{-/-}$ and wild-type mice did not reveal any obvious differences when staining with antibodies against common histone modifications such as H3K9me3 (unpublished data). During spermatogenesis, global sex-specific changes to the epigenome occur due to a wave of DNA demethylation, followed by DNA methylation and chromatin modifications (Figure 4) (Sassone-Corsi, 2002). Because DNA methylation plays a central role in the differentiation of male gamets, its potential influence on histone methyltransferases and histone demethylases may have far-reaching consequences. Indeed, several histone methyltransferases, such as Lysine 9 histone methyltransferase (Suv39h) and Meisetz, have been shown to be important

regulators of germ cell transcription (Hayashi *et al.*, 2005; Peters *et al.*, 2001). It is possible that Alkbh1 regulates the activity of regulatory proteins during spermatogenesis by demethylating histones in target genes.

4.2 Tzfp

Tzfp is a transcription factor belonging to the BTB/POZ-ZF family (Lin *et al.*, 1999). Despite many studies, the exact role of Tzfp is still largely unknown, although a function in lymphocyte activation has been described (Kang *et al.*, 2005; Piazza *et al.*, 2004). In addition, TZFP has been found to be a repressor of GATA (Miaw *et al.*, 2000) and to compose a heterodimer with the FA group C protein (FANCC) (Hoatlin *et al.*, 1999). However, we believe that Tzfp also has a role in spermatogenesis, as the *Tzfp* gene is highly expressed in testis, particularly in pachytene spermatocytes and, to a lesser extent, in round spermatids (Paper III). Using *Tzfp* deficient mice, we found that Tzfp is dispensable for the formation of male germ cells, but that disruption of the gene leads to sex ratio distortion and increases androgen receptor signaling in the Sertoli cells in testes. The lack of a serious phenotype in the absence of Tzfp confirms the findings of two other groups who have generated *Tzfp* deficient mouse strains (Kang *et al.*, 2005; Piazza *et al.*, 2004), both of which report normal development and lack of any gross pathological abnormalities. The other groups, however, did not observe a sex ratio distortion. This discrepancy might be due to differences in genetic background and targeting strategy.

We speculate that Tzfp is part of the transcription program that regulates the production of male gamets. Two of the hallmarks of spermatogenesis and spermiogenesis are the radical changes in chromatin structure and the widespread exchange of most canonical histones for testis specific variants and protamines (Figure 4) (Ward and Coffey, 1991). In addition, a wave of transcription occurs in the post-meiotic male germ cells (Monesi, 1965; Moore, 1971). This is achieved through assembly of transcription complexes and specialized chromatin remodeling enzymes. High transcription activity is required at this stage due to transient silencing of the chromatin upon histone exchange with protamines later in the differentiation process (Sassone-Corsi, 2002). Is it possible that Tzfp plays a part in this complex process, fine-tuning the expression of genes needed to complete the formation of mature spermatozoa? It has previously been shown that TZFP can bind DNA and that it co-localizes with Fanconi Anemia group C protein (FANCC) in nuclear foci (Hoatlin *et al.*, 1999), suggesting a function linked to a transcriptional repression pathway involved in chromatin remodeling.

4.2.1 Tzfp and cell faith determination

TZFP forms a heterodimer with the FANCC (Hoatlin *et al.*, 1999), which is one of at least 13 distinct complementation proteins involved in the autosomal recessive disorder Fanconi anemia (FA) (Moldovan and D'Andrea, 2009). Fanconi anemia is a polygenic trait hypothesized to be a DNA damage repair disease (de Winter *et al.*, 2000). Mice with targeted disruption of the *Fancc* locus have greatly reduced germ-cell numbers by 12.5 dpc due to reduced mitotic proliferation of primordial germ cells, and adult animals display reduced fertility (Nadler and Braun, 2000; Wong *et al.*, 2003). This phenotype suggests that the Fancc protein functions prior to meiosis and that it is required for mitotic proliferation of the PGCs.

In addition to strict transcriptional- and translational regulation during pre-and post mitotic stages of spermatogenesis, transcriptional regulation also plays an important role early in the mitotic phase of spermatogenesis, when spermatogonial cells face the choice of self-renewal or differentiation. One of the players in this decision-making process is Promyelocytic leukemia zinc-finger (Plzf), a BTB/POZ-ZF protein highly related to Tzfp (Costoya and Pandolfi, 2001). The BTB/POZ domain recruits Polycomb proteins, leading to HDAC recruitment and ultimately a stable and heritable repression of specific developmental genes (Barna *et al.*, 2002). Plzf is encoded by the *Zfp145* gene, and *Zfp145*-null mice show progressive loss of spermatogonial cells and increased apoptosis (Buaas *et al.*, 2004; Costoya *et al.*, 2004). It is therefore possible that Plzf is important for defining spermatogonial cell fate.

The striking homology between PLZF and TZFP suggests that the two proteins have similar functions. PLZF has been shown to bind to specific DNA target sequences (Li *et al.*, 1997), and TZFP seems to bind to the same targets (Hoatlin *et al.*, 1999). Despite the high degree of homology, the two proteins display different expression patterns. *Zfp145* expression is restricted to gonocytes and undifferentiated spermatogonia (Costoya *et al.*, 2004), whereas *Tzfp* is expressed in pachytene spermatocytes and round spermatids (Paper III). Combined, these results indicate that Tzfp and Plzf regulate cell faith determination in association with Fancc at different stages in spermatogenesis.

4.2.2 Tzfp regulates gene expression

In 2001, Tang and colleagues described the regulatory role of Tzfp on the *Aie1* gene, which encodes the Aurora-C kinase protein (Tang *et al.*, 2001). Via its zinc fingers, Tzfp binds to a specific DNA sequence, repressing the target gene. This sequence, TGTACAGTGT, called the Tzfp binding sequence (tbs), is located at the upstream flanking sequence of the *Aie1* gene. In accordance with the findings of Tang *et al.*, the expression of *Aie1* is elevated in $Tzfp^{GTi/GTi}$ testes (Paper III), confirming that Tzfp has a repressive effect on the *Aie1* gene. The repressive effect of this Tzfp-DNA interaction is possibly due to the recruitment of N-

CoR/SMRT complexes by the BTB/POZ domain located in the N-terminal of Tzfp, as is found for the human B cell lymphoma (BCL-6) and PLZF oncoproteins (Dhordain *et al.*, 1997; Lin *et al.*, 1998; Wong and Privalsky, 1998). The N-CoR/SMRT complex contains HADAC3 (Guenther et al., 2000; Li et al., 2000), which, interestingly, has GATA as one of its substrates (Ozawa *et al.*, 2001; Yang and Seto, 2007). As described above, TZFP is a known repressor of GATA (Miaw *et al.*, 2000; Tsuzuki and Enver, 2002).

Tzfp probably exists in two isoforms, both of which contain the N-terminal BTB/POZ domain (Lin *et al.*, 1999; Miaw *et al.*, 2000). Only the longer variant, however, which is expressed in testes, contain the C-terminal zinc fingers. This means that only the longer testis specific variant is likely to target genes containing the Tzfp binding sequence. It is therefore likely that the isoform found in testes has a different activity from that found in lymphocytes.

Aurora-C is a testis specific kinase belonging to the Aurora kinase family (Tseng *et al.*, 1998). The Aurora proteins are a conserved family of serine/ threonine kinases important for cell division (Andrews *et al.*, 2003; Nigg, 2001; Carmena and Earnshaw, 2003). Aurora-A and -B are known to play distinct roles in mitosis, whereas Aurora-C is predominantly expressed in meiotically dividing spermatocytes (Bernard *et al.*, 1998; Tang *et al.*, 2006; Tseng *et al.*, 1998), where it associates with INCENP (Tang *et al.*, 2006). Studies on mouse oocytes have shown that inhibition of Aurora-C induces abnormal kinetochore–microtubule attachment, premature chromosome separation, and cytokinesis failure in meiosis I (Yang *et al.*, 2010). Recently, it was reported that overexpression of Aurora-C in cell lines induces abnormal cell division resulting in centrosome amplification and multinucleation (Khan *et al.*, 2011). Such morphological abnormalities were not observed in the $Tzfp^{GTi/GTi}$ samples (Paper III), indicating that the *Aie1* overexpression is not dramatic enough to cause problems with meiotic divisions. Combined, these studies show that correct expression of the Aurora-C protein is important for cytokinesis during meiosis, and that Tzfp is one of perhaps several factors that ensure correct expression levels.

As described earlier, post-translational modification of histone amino acid residues can result in recruitment of effector proteins that may transiently affect the transcription rate of target genes (Bartke *et al.*, 2010; Kouzarides, 2007). Additional modifications of nearby residues may lead to changes in the binding ability of the protein, modifying its effect on the transcription rate (Fischle *et al.*, 2003). The testis contains many specific effector proteins, and also many testis specific histone variants. Combined, this gives great variety and plasticity to the male germline chromatin. The presence of specific modification sites on tH2B and other testis specific histones can thus generate many combinations of modifications e.g. acetylation and/or methylation that could uniquely affect transcription of various loci. The modification sites could be targeted by HDACs and germ cell kinases, such as Aurora-C,

both possibly regulated by Tzfp, generating unique patterns of acetylation and/or methylation. These patterns may be important for fine-tuning the massive transcription program during spermatogenesis.

4.2.3 Loss of Tzfp leads to increased Androgen Receptor signaling

The Androgen Receptor (AR) is a member of the nuclear receptor superfamily and can be detected in Sertoli cells (SC), the Peritubular myoid (PM) cells, and cells in the interstitial spaces of the testis (Kimura *et al.*, 1993; Bremner *et al.*, 1994). The AR functions as a ligand-dependent transcription factor and regulates the expression of an array of androgen-responsive genes by binding directly to androgen response elements (AREs) (Eder *et al.*, 2003; Jiang and Wang, 2003). AR and androgens are vital for spermatogonial differentiation, the meiotic progression of spermatocytes, for the maturation of spermatids into elongated spermatids, and for spermiation (Maclean and Wilkinson, 2005; Verhoeven, 1992).

In $Tzfp^{GTi/GTi}$ testes, the level of *AR* transcript was found to be similar to wild-type levels, indicating that Tzfp has no effect on *AR* transcription (Paper III). However, when staining wild-type and $Tzfp^{GTi/GTi}$ testis sections with anti-AR, a stronger signal was detected in the nucleus of Sertoli cells in the $Tzfp^{GTi/GTi}$ testes compared to the wild-type. This indicates that AR localization, turnover and/or activity are altered upon Tzfp removal. The amount of AR peaks in stage VII-VIII tubuli, indicating a crucial role for AR signaling in the cell types found in these tubules (Bremner *et al.*, 1994). Stage VII-VIII tubuli contain *Tzfp*-expressing mid-pachytene spermatocytes, and step 7-8 spermatids. Tzfp and AR are not found in the same cell types within the testes, however, intricate interactions between Sertoli and germ cells via desmosome-like junctions, ectoplasmic specializations and gap junctions ensure communication between the cell types and are important for proper germ cell differentiation (Mruk and Cheng, 2004).

Like most steroid hormone receptors, AR is characterized by a conserved structural and functional organization (Kumar *et al.*, 2004; Mangelsdorf *et al.*, 1995) and contains a ligand-independent transactivation domain, a highly conserved DNA-binding domain, a homo- and heterodimerization domain, and a ligand-binding domain that harbours a ligand-inducible transactivation function , all of which are subjected to post transcriptional modifications (Faus and Haendler, 2006). These modifications serve to regulate interactions essential for transactivation, and to regulate the receptor turnover. Interestingly, AR activity is specifically down-regulated by HDAC1, and is a direct target for Mdm2-mediated ubiquitylation and proteolysis (Gaughan *et al.*, 2002; Gaughan *et al.*, 2005) In addition, HDAC4 can bind and inhibit the activity of AR (Yang *et al.*, 2011), a process found to be dependent on small ubiquitin-like modifier (SUMO)ylation of the AR. Knock-down of HDAC4 increases the activity of endogenous AR. These results suggest that acetylation and deacetylation is an

important mechanism for AR regulation (Gaughan *et al.*, 2005). It is thus tempting to speculate that lack of Tzfp leads to reduced levels of recruited HDAC, increasing the levels of AR due to reduced levels of ubiquitylation and SUMOylation.

GATA-1 is another candidate for a protein under Tzfp control that may regulate AR activity. As described above, Tzfp is known to repress GATA factors (Miaw *et al.*, 2000; Tsuzuki and Enver, 2002), and in addition, *Gata-1* expression was increased in the $Tzfp^{GTi/GTi}$ testes (Paper III). GATA-1 exerts a temporal expression in Sertoli cells (Viger *et al.*, 2004; Ketola *et al.*, 1999) and is only found in stage VI–XII tubules of the seminiferous cycle (Yomogida *et al.*, 1994; Ketola *et al.*, 2002). This cyclic expression appears to be dependent on the presence of maturing germ cells, and coincides with the androgen-sensitive stages VII–VIII (Bremner *et al.*, 1994) and the stages containing cells with the highest *Tzfp* expression (Paper III). GATA-1 interacts with several HDAC proteins (Watamoto *et al.*, 2003), of which HDAC1 has been shown to bind AR and specifically down-regulate AR transcriptional activity (Gaughan *et al.*, 2002). GATA-1 also regulates the androgen receptor corepressor 19kDa (ARR19), which co-translocates into the nucleus with AR and suppresses AR activity through the recruitment of HDAC4 (Qamar *et al.*, 2009; Jeong *et al.*, 2004).

4.2.4 Alkbh1 is a possible Tzfp interaction partner

Our group identified ALKBH1 as a potential interaction partner of TZFP (Paper II). The ALKBH1-TZFP interaction was first identified by screening a testis library by yeast two-hybrid (Y2H) analysis, and was further verified by Dot blot analysis. The Y2H screen yielded four positive clones, of which three represented the *TZFP* gene. Although we have not been able to verify the interaction through Co-IP analysis, the results indicate that these two proteins may be partners *in vivo*.

TZFP has been found to participate in the regulation of osteoblastic differentiation via the BMP2 signaling pathway (Ikeda *et al.*, 2007). This is highly interesting, as several BMPs were found to have altered expression in the $Alkbh1^{-/-}$ developing embryos, probably contributing to the phenotypes of eye and skeletal defects observed in these animals (Paper I). Ikeda and co-workers showed that *TZFP* expression in cells is upregulated upon BMP2 induction, and that this increases the expression of osteoblastic differentiation markers (Ikeda *et al.*, 2007). No skeletal malformation phenotype was observed in the Tzfp deficient mice (Paper III) (Piazza *et al.*, 2004), indicating functional redundancy. However, combined with our results from the *Alkbh1* deficient mice, these results suggest that Tzfp and Alkbh1 may be linked to a transcriptional regulation pathway involved in BMP2-regulated differentiation. These observations prompted me to generated a *Tzfp/Alkbh1* double knockout to investigate wether this would create a more severe phenotype. The mice, however, had reduced fertility and viability comparable to the single $Alkbh1^{-/-}$ knockout mouse (unpublished data). This

suggests an epistatic relationship between the two proteins, i.e. that removal of the second protein does not result in an additive effect on the observed phenotype.

As described above, Mrj1 has been suggested as an Alkbh1 interaction partner in placenta (Pan *et al.*, 2008). Thus, Alkbh1 may have several interaction partners in different tissues or stages of development, leading to the diverse phenotype observed in *Alkbh1*$^{-/-}$ mice. Mrj can, like Tzfp, potentially lead to gene silencing through HDAC recruitment. Alkbh1-Mrj interaction in trophoblast lineages seems to be necessary for proper development of the placenta (Pan *et al.*, 2008), where binding of Alkbh1 to Mrj inhibits HDAC recruitment, leading to gene regulation. In other tissues like the testis, Alkbh1 may bind Tzfp, reducing its ability to recruit HDACs to target genes. The Tzfp target genes probably include (but may not be restricted to) genes that contain the tbs sequence (Tang *et al.*, 2001).

Interestingly, several HDACs can associate with different GATA factors (Ozawa *et al.*, 2001; Watamoto *et al.*, 2003), suggesting a mechanism where the transcription of Alkbh1-Tzfp-targeted genes are modified through alterations in HDAC and GATA recruitment. Tzfp is known to be a repressor of GATA (Miaw *et al.*, 2000; Tsuzuki and Enver, 2002), a repression that probably occurs through HDAC recruitment. We found *Gata-1* to be upregulated in the *Tzfp*$^{GTi/GTi}$ testes (Paper III), indicating that Tzfp indeed has a repressive effect on members of the GATA family. No effect was seen on the protein level, indicating that other mechanisms are involved in ensuring sufficient amounts of GATA-1 during spermatogenesis. It is uncertain how the repression of HDACs by Alkbh1-Tzfp complexes can mediate such diverse biological functions as suggested in the *Alkbh1*$^{-/-}$ animals. This indicates that in addition to GATA factors, additional transcription factors are also likely subjected to HDAC regulation through direct or indirect interactions.

The target genes of Alkbh1 and Tzfp are still largely unknown, although we and other groups have identified some potential targets, such as the 1082B piRNA cluster in male germ cells (Paper II).

4.3 piRNAs and their role in epigenetics and transposon silencing

The PIWI pathway acts via post-transcriptional and transcriptional mechanisms to silence retrotransposons in prenatal and prepubertal mouse testes (Aravin *et al.*, 2007a; Aravin *et al.*, 2007b; Kuramochi-Miyagawa *et al.*, 2008). This provides an intriguing model of RNA-directed epigenetic control in mammals. The mouse genome has three members of the *Piwi* sub-family; *Miwi* (Mouse Piwi), *Mili* (Mouse Piwi-like), and *Miwi2* (Peters and Meister, 2007), all of which are predominantly expressed in the male germline. Only *Mili* is detected in the female germline, but loss of this protein does not affect female fertility (Aravin *et al.*,

2008; Kuramochi-Miyagawa *et al.*, 2001; Watanabe *et al.*, 2008). In males, however, loss of any of the three proteins causes spermatogenic arrest (Carmell *et al.*, 2007; Deng and Lin, 2002; Kuramochi-Miyagawa *et al.*, 2004).

The association of PIWI proteins to piRNAs is well described (Aravin *et al.*, 2006; Grivna *et al.*, 2006a; Houwing *et al.*, 2007; Saito *et al.*, 2006). The exact biogenesis and function of piRNAs are, however, yet relatively poorly understood, although a function in transposon silencing has been established in both *Drosophila* and mice (Aravin *et al.*, 2007b; Aravin *et al.*, 2008; Carmell *et al.*, 2007; Brennecke *et al.*, 2007). A significant subpopulation of the piRNAs in *Drosophila*, mice, and zebrafish are derived from intergenic regions containing transposable elements and repetitive sequences, and many PIWI proteins have been found to be required for transposon repression in the germline (Brennecke *et al.*, 2007; Houwing *et al.*, 2007; Kalmykova *et al.*, 2005). High-throughput sequencing offer new insight in the different subpopulations and distributions of piRNAs, and we and others have shown that an overwhelming amount of piRNAs are present in mouse testes (Gan *et al.*, 2011) (Paper II). We found that piRNAs present in pachytene spermatocytes have a peak length of 29-30 nt with a preference for a 5'U, indicating that these piRNAs are mainly primary piRNAs interacting with Miwi, produced through processing of primary transcripts. These analyses also show that the majority of piRNAs from adult animals, unlike piRNAs from prenatal and prepubertal mice, do not map to repeat-associated regions of the genome, such as transposons. Rather, they map to piRNA clusters in repeat-deprived intergenic regions. Moreover, a small fraction of pachytene piRNAs map to exons, introns, and 5' and 3' UTRs, indicating that a subpopulation of piRNAs may be involved in transcription or translation regulation. A large fraction of piRNAs in prenatal and prepubertal testes also map to non-repeat associated regions (Gan *et al.*, 2011; Ma *et al.*, 2009), indicating that the PIWI-piRNA pathway might play additional roles to transposon control in these animals.

4.3.1 The role of Tzfp and Alkbh1 in piRNA regulation

The organization of intergenic piRNAs into clusters and the processing of piRNAs from long primary transcripts, are relatively well described (Aravin *et al.*, 2006; Brennecke *et al.*, 2007; Ragan *et al.*, 2012). How these clusters are regulated and how the primary transcripts are processed into individual piRNAs, are, however, unknown. Until now, the detection of piRNA clusters has relied on bioinformatics rather than the detection and sequencing of primary piRNA transcripts. This is probably the reason why cluster 1082B, which we identified to be regulated by Alkbh1 and Tzfp, was initially annotated as part of a large cluster called 1082 (http://pirnabank.ibab.ac.in). Based on our sequencing data, this cluster was divided into three, called 1082A-C, each most likely giving rise to one primary transcript (Paper II). Only the middle cluster is regulated by Tzfp and Alkbh1, as shown by the dramatic

upregulation of individual piRNAs originating from this region in Tzfp and Alkbh1 deficient mice. These findings indicate that at least a subset of the intergenic piRNA clusters are individually regulated by transcription factors and other regulatory proteins.

Mouse testes contain many long ESTs and pseudogenes with unknown functions, some of which, like the *4933440M02Rik*, seem to be the origin of piRNAs (Gan *et al.*, 2011). This indicates that some of these transcripts may in fact be piRNA precursors. Given the high expression of several C_2H_2-type zinc finger transcription factors in testes, it is tempting to speculate that piRNA precursor transcripts are individually regulated by specific proteins, like the C_2H_2-containing transcription factor Tzfp, determining the timing and amount of piRNA expression.

The role of non-repeat pachytene piRNAs in adult testes is a mystery. We found that the massive upregulation of piRNA cluster 1082B seen in the *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ mice is followed by a downregulation of L1 and IAP transcripts (Paper II). The piRNAs in this cluster are derived from non-repetitive regions in the genome and are not complementary to the transposon transcripts, indicating that they act through an unknown mechanism. Because the endonuclease activity of Mili and Miwi is dependent upon high degree of complementarity between piRNA and target transcript (Reuter *et al.*, 2011), these piRNAs probably work through a mechanism that does not involve cleavage of the target transcript. The piRNAs derived from cluster 1082B could possibly work through a miRNA-like mechanism that only requires partial complementarity and lead to translational repression rather than cleavage. Recently, however, Vourekas and co-workers found no evidence of the existence of so-called seed regions in pachytene piRNAs, indicating that in the postnatal male germ line, Mili and Miwi do not use non-repeat derived piRNAs as direct guides to target transcripts (Vourekas *et al.*, 2012).

In *Drosophila*, Piwi proteins are required for the clearing of maternal mRNAs in preparation for the maternal-to-zygotic transition (MZT) (Rouget *et al.*, 2010). Here, PIWI-piRNA complexes recruit the deadenylation machinery to the 3'UTR of target transcripts, destabilizing the mRNA. It is possible that the piRNAs derived form cluster 1082B work through a similar mechanism, leading to degradation of transposon mRNAs. More work is needed to identify the exact mechanism through which non-repeat piRNAs work in adult mouse testes.

Initially, we found the downregulation of L1 and IAP in *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ mice somewhat surprising, as active TEs pose a serious threat to genome integrity. Successful expansion by selfish genetic elements can only occur if increased copy numbers can be transmitted to the next generation. Why would piRNAs that could potentially silence active

transposons normally be repressed? The lack of complete repression of L1 elements in germ cells could raise the question whether transposon activity may have a beneficial role in the host. Studies performed on whole genomes suggest that over one-third of mammalian genomes are the result, directly or indirectly, of L1 retrotranspositioning (Han and Boeke, 2005). It is thus possible that transposons drive evolution by increasing genetic diversity. However, a closely regulated control mechanism must be in place to avoid detrimental damage to the genome integrity by active transposons. It thus seems like a well established balance between increased mutation rate driving evolution and TE silencing mechanisms have evolved in the mouse germline.

4.4 Alkbh1 and Tzfp mutant mice – differences and similarities

The identification of Alkbh1 and Tzfp as possible interaction partners, and the similar expression profiles of the genes encoding the proteins in testes, suggests that the two proteins are involved in the same regulatory mechanisms. The phenotypes displayed in *Alkbh1*$^{-/-}$and *Tzfp*$^{GTi/GTi}$ mice, however, differ considerably (Paper I; Paper III), indicating that they also are involved in different pathways. The *Alkbh1*$^{-/-}$ mouse displays serious developmental defects, whereas no such phenotype is observed in *Tzfp*$^{GTi/GTi}$ mice. In addition, the defects observed in testes also vary between the two mouse Alkbh1 and Tzfp deficient mice.

The different phenotypes can, at least in part, be explained by the different expression profiles of the two genes during embryonic development and in somatic tissues. The *Alkbh1* transcript is detected in most adult tissues, and is also relatively highly expressed during embryonic development (Paper I). This broad expression pattern indicates that Alkbh1 has a role in many organs and during embryonic development. Additionally, Alkbh1 seems to have different interaction partners, and the relatively complex phenotype observed in the *Alkbh1*$^{-/-}$ mice is thus probably a result of dysregulation and dysfunction of the different Alkbh1 interaction partners.

Although a role for Tzfp in hematopoetic cells has been established (Kang *et al.*, 2005; Miaw *et al.*, 2000), the *Tzfp* gene is expressed almost exclusively in the testis of adult animals, indicating that the function of Tzfp is rather testis specific (Paper III). In *Alkbh1*$^{-/-}$ mice, the number of apoptotic cells increased with age, with apoptotic spermatocytes and spermatids seen in a fraction of the tubules in adult animals (Paper I). This defect was not observed in *Tzfp*$^{GTi/GTi}$ mice. In fact, adult *Tzfp*$^{GTi/GTi}$ testes were found to have a reduced number of apoptotic cells when compared to wild-type (Paper III). I therefore speculate that lack of Alkbh1 results in an increased number of dysfunctional spermatocytes and spermatids, whereas removal of Tzfp disrupts cell faith determination leading to decreased removal of

redundant and/or damaged spermatocytes. Both proteins, however, seem to work through epigenetic regulatory mechanisms, possibly by leading to deacetylation and demethylation of histones in the germline chromatin. The lack of a severe phenotype in *Tzfp*$^{GTi/GTi}$ mice indicates redundancy in the biological pathways in which the protein is involved.

Perhaps the most striking difference between the two mice mutants is the presence of sex ratio distortion with opposite skewing. In *Alkbh1*$^{-/-}$ mice, one female pup is born for every three to four males ($p=1.4x10^{-6}$ using χ^2 test) (Paper I), whereas fewer males than expected are born with *Tzfp*$^{GTi/GTi}$ genotype ($p=0.003$ using χ^2 test) (Paper III). Both models thus seem to lead to a transmission ratio distortion (TRD). TRD involves a deviation from the expected Mendelian genotypic frequencies due to a non-Mendelian inheritance pattern for a given genetic variant (Lyon, 2003). In females, TRD can occur by preferential entry of one allele into the polar body at meiosis, whereas in males, it occurs if one allele impairs either the formation or the function of sperm. One example of mammalian TDR is Robertsonian translocations involving mouse chromosome 6, where male mice heterozygous for the Rb(6.16) translocation show a transmission distortion of the segregants in the resulting zygotes (Aranha and Martin-DeLeon, 1995). These mice exhibit sex ratio skewing in favor of male offspring, indicating that there must be some functional asymmetry between X and Y spermatozoa. The different outcomes of the transmission rate distortion of the *Alkbh1* and the *Tzfp* mutant alleles indicate that they lead to the involvement of different distorter and responder genes which collaboratively lead to TRD and sex ratio distortions in favor of male and female pups, respectively.

5 Conclusions and Perspectives

The AlkB enzyme in *E. coli* repairs DNA lesions by a well-described mechanism (Trewick *et al.*, 2002; Falnes *et al.*, 2002). The biological role of many of the AlkB homologs in mice and humans has, however, for many years been a mystery. The role of ALKBH1, the protein with the highest degree of homology to *E. coli* AlkB, is still relatively unclear. As described in the introduction of this thesis, several somewhat contradictory studies have been published, suggesting different functions for the ALKBH1 protein (Muller *et al.*, 2010; Pan *et al.*, 2008; Westbye *et al.*, 2008). Recent studies suggest that ALKBH1 may have histones as its substrate rather than DNA, and might work through a mechanism similar to that of histone demethylases containing the JmjC-domain (Tsukada *et al.*, 2006) (Ougland *et al.*, in press). Studies on *Alkbh1* deficient mice indicate that the protein has an important role in cell differentiation and development (Paper I). Taken together, the results from our and previous studies indicate that Alkbh1 plays a role in ensuring proper gene expression during differentiation and development. More research is needed to pinpoint the exact mechanism through which Alkbh1 works. This includes further studies on the role of Alkbh1 in spermatogenesis. Identification of more Alkbh1 interaction partners also awaits closer examination.

One possible ALKBH1 interaction partner, TZFP, was identified by our group (Paper II). The murine homolog encoding this protein is highly expressed in pachytene spermatocytes and encodes a BTB/POZ-ZF protein with a mostly unknown function. Studies using *Tzfp* deficient mice revealed that lack of this protein leads to increased Androgen Receptor signaling in Sertoli cells (Paper III), but the mechanism through which this protein works remains unknown. A possible mechanism involves binding of the Tzfp protein to genes that contain the tbs sequence (Tang *et al.*, 2001) and recruitment of N-CoR/SMRT complex by the BTB/POZ domain (Dhordain *et al.*, 1997), leading to gene silencing by HDAC recruitment. Further studies are needed to determine the exact role of Tzfp and Alkbh1 in spermatogenesis. This includes verification of the Alkbh1-Tzfp interaction and determination of target genes.

We identified the piRNA cluster 1082B on chromosome 7 in the mouse genome as a target sequence for Alkbh1 and Tzfp (Paper II). This links the function of these two proteins to the PIWI-piRNA pathway. Studies of the PIWI-piRNA pathway have revealed diverse biological and molecular functions; including roles in germline specification, gametogenesis and stem cell maintenance, epigenetic programming, transposon silencing, genome protection, position effect variegation, hybrid dysgenesis, genomic imprinting, and posttranscriptional regulation of mRNAs.

One of the most well-described functions of piRNAs is that of TE repression in the germline, both at the transcriptional and posttranscriptional level (Siomi *et al.*, 2011). However, despite intensive research the past decade, the molecular mechanisms underlying these processes remain elusive. This includes the production of primary piRNAs, such as the transcriptional control and processing of the primary piRNA precursors. Our research has brought us one step closer in the quest for understanding this process, as we provide evidence for an individually regulated piRNA cluster by the proteins Alkbh1 and Tzfp in mouse (Paper II). In addition, the roles of non-repeat pachytene piRNAs in adult testes have been largely unknown. However, we show that at least a sup-population of these abundant piRNAs may be involved in transposon silencing. Further research is needed to study this in more detail. Also, how desilencing of transposons are connected to the *Piwi* mutant phenotypes, such as loss of germline stem cells and embryonic patterning defects in *Drosophila* (Cox *et al.*, 2000), germline stem cell loss and spermatogenic arrest in mice (Carmell *et al.*, 2007; Deng and Lin, 2002; Kuramochi-Miyagawa *et al.*, 2004), and inhibition of stem cell function in lower organisms such as worms (Funayama *et al.*, 2010), remains to be explained. Another role with many unanswered questions is the PIWI-piRNA function in epigenetic control. HP1 has been found to be involved in heterochromatin formation with PIWI proteins (Brower-Toland *et al.*, 2007), but the relationship between histone methyltransferases, DNA methyltransferases, and PIWI /piRNA is not well understood.

References

Aas,P.A., Otterlei,M., Falnes,P.O., Vagbo,C.B., Skorpen,F., Akbari,M., Sundheim,O., Bjoras,M., Slupphaug,G., Seeberg,E., *et al.* (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature *421*, 859-863.

Ahuja,N., Li,Q., Mohan,A.L., Baylin,S.B., and Issa,J.P. (1998). Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res. *58*, 5489-5494.

Andrews,P.D., Knatko,E., Moore,W.J., and Swedlow,J.R. (2003). Mitotic mechanics: the auroras come into view. Curr. Opin. Cell Biol. *15*, 672-683.

Aranha,I.P. and Martin-DeLeon,P.A. (1995). Mouse chromosome 6 in Rb translocations: consequences in singly and doubly heterozygous males. Cytogenet. Cell Genet. *69*, 253-259.

Aravin,A., Gaidatzis,D., Pfeffer,S., Lagos-Quintana,M., Landgraf,P., Iovino,N., Morris,P., Brownstein,M.J., Kuramochi-Miyagawa,S., Nakano,T., *et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin,A.A., Hannon,G.J., and Brennecke,J. (2007a). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761-764.

Aravin,A.A., Naumova,N.M., Tulin,A.V., Vagin,V.V., Rozovsky,Y.M., and Gvozdev,V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr. Biol. *11*, 1017-1027.

Aravin,A.A., Sachidanandam,R., Bourc'his,D., Schaefer,C., Pezic,D., Toth,K.F., Bestor,T., and Hannon,G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell *31*, 785-799.

Aravin,A.A., Sachidanandam,R., Girard,A., Fejes-Toth,K., and Hannon,G.J. (2007b). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Aravin,A.A., van der Heijden,G.W., Castaneda,J., Vagin,V.V., Hannon,G.J., and Bortvin,A. (2009). Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS. Genet. *5*, e1000764.

Aravind,L. and Koonin,E.V. (2001). The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol. *2*, RESEARCH0007.

Auerbach,A.D. (1988). A test for Fanconi's anemia. Blood *72*, 366-367.

Baarends,W.M., Wassenaar,E., van der,L.R., Hoogerbrugge,J., Sleddens-Linkels,E., Hoeijmakers,J.H., de,B.P., and Grootegoed,J.A. (2005). Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. Mol. Cell Biol. *25*, 1041-1053.

Bandyopadhyay,A., Tsuji,K., Cox,K., Harfe,B.D., Rosen,V., and Tabin,C.J. (2006). Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis. PLoS. Genet. *2*, e216.

Bannister,A.J. and Kouzarides,T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381-395.

Bannister,A.J., Schneider,R., Myers,F.A., Thorne,A.W., Crane-Robinson,C., and Kouzarides,T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J. Biol. Chem. *280*, 17732-17736.

Bannister,A.J., Zegerman,P., Partridge,J.F., Miska,E.A., Thomas,J.O., Allshire,R.C., and Kouzarides,T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Barchi,M., Mahadevaiah,S., Di,G.M., Baudat,F., de Rooij,D.G., Burgoyne,P.S., Jasin,M., and Keeney,S. (2005). Surveillance of different recombination defects in mouse spermatocytes yields distinct responses despite elimination at an identical developmental stage. Mol. Cell Biol. *25*, 7203-7215.

Bardwell,V.J. and Treisman,R. (1994). The POZ domain: a conserved protein-protein interaction motif. Genes Dev. *8*, 1664-1677.

Barna,M., Hawe,N., Niswander,L., and Pandolfi,P.P. (2000). Plzf regulates limb and axial skeletal patterning. Nat. Genet. *25*, 166-172.

Barna,M., Merghoub,T., Costoya,J.A., Ruggero,D., Branford,M., Bergia,A., Samori,B., and Pandolfi,P.P. (2002). Plzf mediates transcriptional repression of HoxD gene expression through chromatin remodeling. Dev. Cell *3*, 499-510.

Bartel,D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bartke,A. (1995). Apoptosis of male germ cells, a generalized or a cell type-specific phenomenon? Endocrinology *136*, 3-4.

Bartke,T., Vermeulen,M., Xhemalce,B., Robson,S.C., Mann,M., and Kouzarides,T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell *143*, 470-484.

Berger,S.L., Kouzarides,T., Shiekhattar,R., and Shilatifard,A. (2009). An operational definition of epigenetics. Genes Dev. *23*, 781-783.

Bernard,M., Sanseau,P., Henry,C., Couturier,A., and Prigent,C. (1998). Cloning of STK13, a third human protein kinase related to Drosophila aurora and budding yeast Ipl1 that maps on chromosome 19q13.3-ter. Genomics *53*, 406-409.

Bertone,P., Stolc,V., Royce,T.E., Rozowsky,J.S., Urban,A.E., Zhu,X., Rinn,J.L., Tongprasit,W., Samanta,M., Weissman,S. *et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *306*, 2242-2246.

Betel,D., Sheridan,R., Marks,D.S., and Sander,C. (2007). Computational analysis of mouse piRNA sequence and biogenesis. PLoS. Comput. Biol. *3*, e222.

Beyret,E. and Lin,H. (2011). Pinpointing the expression of piRNAs and function of the PIWI protein subfamily during spermatogenesis in the mouse. Dev. Biol. *355*, 215-226.

Beyret,E., Liu,N., and Lin,H. (2012). piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. Cell Res.

Bird,A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6-21.

Bird,A., Taggart,M., Frommer,M., Miller,O.J., and Macleod,D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Blanco-Rodriguez,J., Martinez-Garcia,C., and Porras,A. (2003). Correlation between DNA synthesis in the second, third and fourth generations of spermatogonia and the occurrence of apoptosis in both spermatogonia and spermatocytes. Reproduction. *126*, 661-668.

Boorstein,R.J., Cummings,A., Jr., Marenstein,D.R., Chan,M.K., Ma,Y., Neubert,T.A., Brown,S.M., and Teebor,G.W. (2001). Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. J. Biol. Chem. *276*, 41991-41997.

Booth,M.J., Branco,M.R., Ficz,G., Oxley,D., Krueger,F., Reik,W., and Balasubramanian,S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science *336*, 934-937.

Bose,J., Gruber,A.D., Helming,L., Schiebe,S., Wegener,I., Hafner,M., Beales,M., Kontgen,F., and Lengeling,A. (2004). The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. J. Biol. *3*, 15.

Botuyan,M.V., Lee,J., Ward,I.M., Kim,J.E., Thompson,J.R., Chen,J., and Mer,G. (2006). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. Cell *127*, 1361-1373.

Bourc'his,D. and Bestor,T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Branciforte,D. and Martin,S.L. (1994). Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. Mol. Cell Biol. *14*, 2584-2592.

Braun,R.E. (2000). Temporal control of protein synthesis during spermatogenesis. Int. J. Androl -94.

Bremner,W.J., Millar,M.R., Sharpe,R.M., and Saunders,P.T. (1994). Immunohistochemical localization of androgen receptors in the rat testis: evidence for stage-dependent expression and regulation by androgens. Endocrinology *135*, 1227-1234.

Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R., and Hannon,G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089-1103.

Brennecke,J., Stark,A., Russell,R.B., and Cohen,S.M. (2005). Principles of microRNA-target recognition. PLoS. Biol. *3*, e85.

Brouha,B., Schustak,J., Badge,R.M., Lutz-Prigge,S., Farley,A.H., Moran,J.V., and Kazazian,H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. U. S. A *100*, 5280-5285.

Brower-Toland,B., Findley,S.D., Jiang,L., Liu,L., Yin,H., Dus,M., Zhou,P., Elgin,S.C., and Lin,H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev. *21*, 2300-2311.

Buaas,F.W., Kirsh,A.L., Sharma,M., McLean,D.J., Morris,J.L., Griswold,M.D., de Rooij,D.G., and Braun,R.E. (2004). Plzf is required in adult male germ cells for stem cell self-renewal. Nat. Genet. *36*, 647-652.

Burgoyne,P.S., Mahadevaiah,S.K., and Turner,J.M. (2009). The consequences of asynapsis for mammalian meiosis. Nat. Rev. Genet. *10*, 207-216.

Burwinkel,B. and Kilimann,M.W. (1998). Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. J. Mol. Biol. *277*, 513-517.

Cantara,W.A., Crain,P.F., Rozenski,J., McCloskey,J.A., Harris,K.A., Zhang,X., Vendeix,F.A., Fabris,D., and Agris,P.F. (2011). The RNA Modification Database, RNAMDB: 2011 update. Nucleic Acids Res. *39*, D195-D201.

Carmell,M.A., Girard,A., van de Kant,H.J., Bourc'his,D., Bestor,T.H., de Rooij,D.G., and Hannon,G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell *12*, 503-514.

Carmena,M. and Earnshaw,W.C. (2003). The cellular geography of aurora kinases. Nat. Rev. Mol. Cell Biol. *4*, 842-854.

Carreau,S. and Hess,R.A. (2010). Oestrogens and spermatogenesis. Philos. Trans. R. Soc. Lond B Biol. Sci. *365*, 1517-1535.

Celeste,A., Petersen,S., Romanienko,P.J., Fernandez-Capetillo,O., Chen,H.T., Sedelnikova,O.A., Reina-San-Martin,B., Coppola,V., Meffre,E., Difilippantonio,M.J., *et al.* (2002). Genomic instability in mice lacking histone H2AX. Science *296*, 922-927.

Cheifetz,S. (1999). BMP receptors in limb and tooth formation. Crit Rev. Oral Biol. Med. *10*, 182-198.

Chen,J.X., Zheng,Y., West,M., and Tang,M.S. (1998). Carcinogens preferentially bind at methylated CpG in the p53 mutational hot spots. Cancer Res. *58*, 2070-2075.

Chuang,J.Z., Zhou,H., Zhu,M., Li,S.H., Li,X.J., and Sung,C.H. (2002). Characterization of a brain-enriched chaperone, MRJ, that inhibits Huntingtin aggregation and toxicity independently. J. Biol. Chem. *277*, 19831-19838.

Chuma,S., Hiyoshi,M., Yamamoto,A., Hosokawa,M., Takamune,K., and Nakatsuji,N. (2003). Mouse Tudor Repeat-1 (MTR-1) is a novel component of chromatoid bodies/nuages in male germ cells and forms a complex with snRNPs. Mech. Dev. *120*, 979-990.

Chuma,S., Hosokawa,M., Tanaka,T., and Nakatsuji,N. (2009). Ultrastructural characterization of spermatogenesis and its evolutionary conservation in the germline: germinal granules in mammals. Mol. Cell Endocrinol. *306*, 17-23.

Coleman,M.L., McDonough,M.A., Hewitson,K.S., Coles,C., Mecinovic,J., Edelmann,M., Cook,K.M., Cockman,M.E., Lancaster,D.E., Kessler,B.M. *et al.* (2007). Asparaginyl hydroxylation of the Notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor. J. Biol. Chem. *282*, 24027-24038.

Cooke,H.J. and Saunders,P.T. (2002). Mouse models of male infertility. Nat. Rev. Genet. *3*, 790-801.

Costa,Y., Speed,R.M., Gautier,P., Semple,C.A., Maratou,K., Turner,J.M., and Cooke,H.J. (2006). Mouse MAELSTROM: the link between meiotic silencing of unsynapsed chromatin and microRNA pathway? Hum. Mol. Genet. *15*, 2324-2334.

Costoya,J.A., Hobbs,R.M., Barna,M., Cattoretti,G., Manova,K., Sukhwani,M., Orwig,K.E., Wolgemuth,D.J., and Pandolfi,P.P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. Nat. Genet. *36*, 653-659.

Costoya,J.A. and Pandolfi,P.P. (2001). The role of promyelocytic leukemia zinc finger and promyelocytic leukemia in leukemogenesis and development. Curr. Opin. Hematol. *8*, 212-217.

Cox,D.N., Chao,A., Baker,J., Chang,L., Qiao,D., and Lin,H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. *12*, 3715-3727.

Cox,D.N., Chao,A., and Lin,H. (2000). piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development *127*, 503-514.

Cuthbert,G.L., Daujat,S., Snowden,A.W., Erdjument-Bromage,H., Hagiwara,T., Yamada,M., Schneider,R., Gregory,P.D., Tempst,P., Bannister, A.J. *et al.* (2004). Histone deimination antagonizes arginine methylation. Cell *118*, 545-553.

Dai,M.S., Chevallier,N., Stone,S., Heinrich,M.C., McConnell,M., Reuter,T., Broxmeyer,H.E., Licht,J.D., Lu,L., and Hoatlin,M.E. (2002). The effects of the Fanconi anemia zinc finger (FAZF) on cell cycle, apoptosis, and proliferation are differentiation stage-specific. J. Biol. Chem. *277*, 26327-26334.

Dai,Y.S., Xu,J., and Molkentin,J.D. (2005). The DnaJ-related factor Mrj interacts with nuclear factor of activated T cells c3 and mediates transcriptional repression through class II histone deacetylase recruitment. Mol. Cell Biol. *25*, 9936-9948.

De Fazio S., Bartonicek,N., Di,G.M., breu-Goodger,C., Sankar,A., Funaya,C., Antony,C., Moreira,P.N., Enright,A.J., and O'Carroll,D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature *480*, 259-263.

de Rooij,D.G. and de Boer P. (2003). Specific arrests of spermatogenesis in genetically modified and mutant mice. Cytogenet. Genome Res. *103*, 267-276.

de Winter,J.P., van der,W.L., de,G.J., Stone,S., Waisfisz,Q., Arwert,F., Scheper,R.J., Kruyt,F.A., Hoatlin,M.E., and Joenje,H. (2000). The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG. Hum. Mol. Genet. *9*, 2665-2674.

Deng,W. and Lin,H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell *2*, 819-830.

Dernburg,A.F., McDonald,K., Moulder,G., Barstead,R., Dresser,M., and Villeneuve,A.M. (1998). Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell *94*, 387-398.

Dhordain,P., Albagli,O., Lin,R.J., Ansieau,S., Quief,S., Leutz,A., Kerckaert,J.P., Evans,R.M., and Leprince,D. (1997). Corepressor SMRT binds the BTB/POZ repressing domain of the LAZ3/BCL6 oncoprotein. Proc. Natl. Acad. Sci. U. S. A *94*, 10762-10767.

DiBello,P.R., Withers,D.A., Bayer,C.A., Fristrom,J.W., and Guild,G.M. (1991). The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers. Genetics *129*, 385-397.

Doege,C.A., Inoue,K., Yamashita,T., Rhee,D.B., Travis,S., Fujita,R., Guarnieri,P., Bhagat,G., Vanti,W.B., Shih,A. *et al.* (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature *488*, 652-655.

Dominissini,D., Moshitch-Moshkovitz,S., Schwartz,S., Salmon-Divon,M., Ungar,L., Osenberg,S., Cesarkas,K., Jacob-Hirsch,J., Amariglio,N., Kupiec,M. *et al.* (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201-206.

Ducy,P. and Karsenty,G. (2000). The family of bone morphogenetic proteins. Kidney Int. *57*, 2207-2214.

Dudley,A.T., Lyons,K.M., and Robertson,E.J. (1995). A requirement for bone morphogenetic protein-7 during development of the mammalian kidney and eye. Genes Dev. *9*, 2795-2807.

Duncan,T., Trewick,S.C., Koivisto,P., Bates,P.A., Lindahl,T., and Sedgwick,B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. Proc. Natl. Acad. Sci. U. S. A *99*, 16660-16665.

Dupressoir,A. and Heidmann,T. (1996). Germ line-specific expression of intracisternal A-particle retrotransposons in transgenic mice. Mol. Cell Biol. *16*, 4495-4503.

Eddy,E.M. (1975). Germ plasm and the differentiation of the germ cell line. Int. Rev. Cytol. *43*, 229-280.

Eder,I.E., Haag,P., Basik,M., Mousses,S., Bektic,J., Bartsch,G., and Klocker,H. (2003). Gene expression changes following androgen receptor elimination in LNCaP prostate cancer cells. Mol. Carcinog. *37*, 181-191.

Falnes,P.O. (2004). Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. Nucleic Acids Res. *32*, 6260-6267.

Falnes,P.O., Johansen,R.F., and Seeberg,E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature *419*, 178-182.

Faus,H. and Haendler,B. (2006). Post-translational modifications of steroid receptors. Biomed. Pharmacother. *60*, 520-528.

Fazi,F., Racanicchi,S., Zardo,G., Starnes,L.M., Mancini,M., Travaglini,L., Diverio,D., Ammatuna,E., Cimino,G., Lo-Coco,F. *et al.* (2007). Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO oncoprotein. Cancer Cell *12*, 457-466.

Fernandez-Capetillo,O., Mahadevaiah,S.K., Celeste,A., Romanienko,P.J., Camerini-Otero,R.D., Bonner,W.M., Manova,K., Burgoyne,P., and Nussenzweig,A. (2003). H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. Dev. Cell *4*, 497-508.

Fischle,W., Wang,Y., and Allis,C.D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Fu,D., Brophy,J.A., Chan,C.T., Atmore,K.A., Begley,U., Paules,R.S., Dedon,P.C., Begley,T.J., and Samson,L.D. (2010). Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. Mol. Cell Biol. *30*, 2449-2459.

Funayama,N., Nakatsukasa,M., Mohri,K., Masuda,Y., and Agata,K. (2010). Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. Evol. Dev. *12*, 275-287.

Gan,H., Lin,X., Zhang,Z., Zhang,W., Liao,S., Wang,L., and Han,C. (2011). piRNA profiling during specific stages of mouse spermatogenesis. RNA. *17*, 1191-1203.

Gandini,L., Lombardo,F., Paoli,D., Caponecchia,L., Familiari,G., Verlengia,C., Dondero,F., and Lenzi,A. (2000). Study of apoptotic DNA fragmentation in human spermatozoa. Hum. Reprod. *15*, 830-839.

Gaughan,L., Logan,I.R., Cook,S., Neal,D.E., and Robson,C.N. (2002). Tip60 and histone deacetylase 1 regulate androgen receptor activity through changes to the acetylation status of the receptor. J. Biol. Chem. *277*, 25904-25913.

Gaughan,L., Logan,I.R., Neal,D.E., and Robson,C.N. (2005). Regulation of androgen receptor and histone deacetylase 1 by Mdm2-mediated ubiquitylation. Nucleic Acids Res. *33*, 13-26.

Gerken,T., Girard,C.A., Tung,Y.C., Webby,C.J., Saudek,V., Hewitson,K.S., Yeo,G.S., McDonough,M.A., Cunliffe,S., McNeill,L.A. *et al.* (2007). The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science *318*, 1469-1472.

Girard,A., Sachidanandam,R., Hannon,G.J., and Carmell,M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Glazier,A.M., Nadeau,J.H., and Aitman,T.J. (2002). Finding genes that underlie complex traits. Science *298*, 2345-2349.

Godt,D., Couderc,J.L., Cramton,S.E., and Laski,F.A. (1993). Pattern formation in the limbs of Drosophila: bric a brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. Development *119*, 799-812.

Grivna,S.T., Beyret,E., Wang,Z., and Lin,H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709-1714.

Grivna,S.T., Pyhtila,B., and Lin,H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc. Natl. Acad. Sci. U. S. A *103*, 13415-13420.

Guenther,M.G., Lane,W.S., Fischle,W., Verdin,E., Lazar,M.A., and Shiekhattar,R. (2000). A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. Genes Dev. *14*, 1048-1057.

Hajkova,P., Erhardt,S., Lane,N., Haaf,T., El-Maarri,O., Reik,W., Walter,J., and Surani,M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. *117*, 15-23.

Halic,M. and Moazed,D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504-516.

Hall,B.K. and Miyake,T. (2000). All for one and one for all: condensations and the initiation of skeletal development. Bioessays *22*, 138-147.

Han,J.S. and Boeke,J.D. (2005). LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? Bioessays *27*, 775-784.

Handel,M.A. (2004). The XY body: a specialized meiotic chromatin domain. Exp. Cell Res. *296*, 57-63.

Harper,J.E., Miceli,S.M., Roberts,R.J., and Manley,J.L. (1990). Sequence specificity of the human mRNA N6-adenosine methylase in vitro. Nucleic Acids Res. *18*, 5735-5741.

Hassa,P.O., Haenni,S.S., Elser,M., and Hottiger,M.O. (2006). Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? Microbiol. Mol. Biol. Rev. *70*, 789-829.

Hayashi,K., Yoshida,K., and Matsui,Y. (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature *438*, 374-378.

He,C. (2010). Grand challenge commentary: RNA epigenetics? Nat. Chem. Biol. *6*, 863-865.

Henriksen,K., Hakovirta,H., and Parvinen,M. (1995). Testosterone inhibits and induces apoptosis in rat seminiferous tubules in a stage-specific manner: in situ quantification in squash preparations after administration of ethane dimethane sulfonate. Endocrinology *136*, 3285-3291.

Hess,R.A. and Renato de,F.L. (2008). Spermatogenesis and cycle of the seminiferous epithelium. Adv. Exp. Med. Biol. *636*, 1-15.

Hirose,Y., Suzuki,R., Ohba,T., Hinohara,Y., Matsuhara,H., Yoshida,M., Itabashi,Y., Murakami,H., and Yamamoto,A. (2011). Chiasmata promote monopolar attachment of sister chromatids and their co-segregation toward the proper pole during meiosis I. PLoS. Genet. *7*, e1001329.

Hoatlin,M.E., Zhi,Y., Ball,H., Silvey,K., Melnick,A., Stone,S., Arai,S., Hawe,N., Owen,G., Zelent,A., and Licht,J.D. (1999). A novel BTB/POZ transcriptional repressor protein interacts with the Fanconi anemia group C protein and PLZF. Blood *94*, 3737-3747.

Houwing,S., Kamminga,L.M., Berezikov,E., Cronembold,D., Girard,A., van den,E.H., Filippov,D.V., Blaser,H., Raz,E., Moens,C.B. *et al.* (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69-82.

Hu,J., Chen,Y.X., Wang,D., Qi,X., Li,T.G., Hao,J., Mishina,Y., Garbers,D.L., and Zhao,G.Q. (2004). Developmental expression and function of Bmp4 in spermatogenesis and in maintaining epididymal integrity. Dev. Biol. *276*, 158-171.

Hunter,P.J., Swanson,B.J., Haendel,M.A., Lyons,G.E., and Cross,J.C. (1999). Mrj encodes a DnaJ-related co-chaperone that is essential for murine placental development. Development *126*, 1247-1258.

Ikeda,R., Yoshida,K., and Inoue,I. (2007). Identification of FAZF as a novel BMP2-induced transcription factor during osteoblastic differentiation. J. Cell Biochem. *101*, 147-154.

Ito,S., D'Alessio,A.C., Taranova,O.V., Hong,K., Sowers,L.C., and Zhang,Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.

Jenkins,C.L. and Raines,R.T. (2002). Insights on the conformational stability of collagen. Nat. Prod. Rep. *19*, 49-59.

Jeong,B.C., Hong,C.Y., Chattopadhyay,S., Park,J.H., Gong,E.Y., Kim,H.J., Chun,S.Y., and Lee,K. (2004). Androgen receptor corepressor-19 kDa (ARR19), a leucine-rich protein that represses the transcriptional activity of androgen receptor through recruitment of histone deacetylase. Mol. Endocrinol. *18*, 13-25.

Jia,G., Fu,Y., Zhao,X., Dai,Q., Zheng,G., Yang,Y., Yi,C., Lindahl,T., Pan,T., Yang,Y.G. *et al.* (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. *7*, 885-887.

Jia,G., Yang,C.G., Yang,S., Jian,X., Yi,C., Zhou,Z., and He,C. (2008). Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. FEBS Lett. *582*, 3313-3319.

Jiang,F. and Wang,Z. (2003). Identification of androgen-responsive genes in the rat ventral prostate by complementary deoxyribonucleic acid subtraction and microarray. Endocrinology *144*, 1257-1265.

Johnston,D.S., Russell,L.D., Friel,P.J., and Griswold,M.D. (2001). Murine germ cells do not require functional androgen receptors to complete spermatogenesis following spermatogonial stem cell transplantation. Endocrinology *142*, 2405-2408.

Jones,C.M., Lyons,K.M., Lapan,P.M., Wright,C.V., and Hogan,B.L. (1992). DVR-4 (bone morphogenetic protein-4) as a posterior-ventralizing factor in Xenopus mesoderm induction. Development *115*, 639-647.

Jones,P.A. (1996). DNA methylation errors and cancer. Cancer Res. *56*, 2463-2467.

Kalmykova,A.I., Klenov,M.S., and Gvozdev,V.A. (2005). Argonaute protein PIWI controls mobilization of retrotransposons in the Drosophila male germline. Nucleic Acids Res. *33*, 2052-2059.

Kaneda,M., Okano,M., Hata,K., Sado,T., Tsujimoto,N., Li,E., and Sasaki,H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature -903.

Kang,B.Y., Miaw,S.C., and Ho,I.C. (2005). ROG negatively regulates T-cell activation but is dispensable for Th-cell differentiation. Mol. Cell Biol. *25*, 554-562.

Karlic,R., Chung,H.R., Lasserre,J., Vlahovicek,K., and Vingron,M. (2010). Histone modification levels are predictive for gene expression. Proc. Natl. Acad. Sci. U. S. A *107*, 2926-2931.

Kataoka,H. and Sekiguchi,M. (1985). Molecular cloning and characterization of the alkB gene of Escherichia coli. Mol. Gen. Genet. *198*, 263-269.

Kawaoka,S., Izumi,N., Katsuma,S., and Tomari,Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. Mol. Cell *43*, 1015-1022.

Ketola,I., Anttonen,M., Vaskivuo,T., Tapanainen,J.S., Toppari,J., and Heikinheimo,M. (2002). Developmental expression and spermatogenic stage specificity of transcription factors GATA-1 and GATA-4 and their cofactors FOG-1 and FOG-2 in the mouse testis. Eur. J. Endocrinol. *147*, 397-406.

Ketola,I., Rahman,N., Toppari,J., Bielinska,M., Porter-Tinge,S.B., Tapanainen,J.S., Huhtaniemi,I.T., Wilson,D.B., and Heikinheimo,M. (1999). Expression and regulation of transcription factors GATA-4 and GATA-6 in developing mouse testis. Endocrinology *140*, 1470-1480.

Khan,J., Ezan,F., Cremet,J.Y., Fautrel,A., Gilot,D., Lambert,M., Benaud,C., Troadec,M.B., and Prigent,C. (2011). Overexpression of active Aurora-C kinase results in cell transformation and tumour formation. PLoS. One. *6*, e26512.

Kimura,N., Mizokami,A., Oonuma,T., Sasano,H., and Nagura,H. (1993). Immunocytochemical localization of androgen receptor with polyclonal antibody in paraffin-embedded human tissues. J. Histochem. Cytochem. *41*, 671-678.

Klattenhoff,C., Bratu,D.P., Ginnis-Schultz,N., Koppetsch,B.S., Cook,H.A., and Theurkauf,W.E. (2007). Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. Dev. Cell *12*, 45-55.

Klug,A. and Schwabe,J.W. (1995). Protein motifs 5. Zinc fingers. FASEB J. *9*, 597-604.

Korvald,H., Falnes,P.O., Laerdahl,J.K., Bjoras,M., and Alseth,I. (2012). The Schizosaccharomyces pombe AlkB homolog Abh1 exhibits AP lyase activity but no demethylase activity. DNA Repair (Amst) *11*, 453-462.

Kotaja,N., Lin,H., Parvinen,M., and Sassone-Corsi,P. (2006). Interplay of PIWI/Argonaute protein MIWI and kinesin KIF17b in chromatoid bodies of male germ cells. J. Cell Sci. *119*, 2819-2825.

Kouzarides,T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kovalovsky,D., Uche,O.U., Eladad,S., Hobbs,R.M., Yi,W., Alonzo,E., Chua,K., Eidson,M., Kim,H.J., Im,J.S., Pandolfi,P.P., and Sant'Angelo,D.B. (2008). The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. Nat. Immunol. *9*, 1055-1064.

Kriaucionis,S. and Heintz,N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Kumar,R., Johnson,B.H., and Thompson,E.B. (2004). Overview of the structural basis for transcription regulation by nuclear hormone receptors. Essays Biochem. *40*, 27-39.

Kuramochi-Miyagawa,S., Kimura,T., Ijiri,T.W., Isobe,T., Asada,N., Fujita,Y., Ikawa,M., Iwai,N., Okabe,M., Deng,W. *et al.* (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Kuramochi-Miyagawa,S., Kimura,T., Yomogida,K., Kuroiwa,A., Tadokoro,Y., Fujita,Y., Sato,M., Matsuda,Y., and Nakano,T. (2001). Two mouse piwi-related genes: miwi and mili. Mech. Dev. *108*, 121-133.

Kuramochi-Miyagawa,S., Watanabe,T., Gotoh,K., Totoki,Y., Toyoda,A., Ikawa,M., Asada,N., Kojima,K., Yamaguchi,Y., Ijiri,T.W. *et al.* (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. *22*, 908-917.

Kurowski,M.A., Bhagwat,A.S., Papaj,G., and Bujnicki,J.M. (2003). Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. BMC. Genomics *4*, 48.

Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. *et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Larsen,F., Gundersen,G., Lopez,R., and Prydz,H. (1992). CpG islands as gene markers in the human genome. Genomics *13*, 1095-1107.

Lau,N.C., Seto,A.G., Kim,J., Kuramochi-Miyagawa,S., Nakano,T., Bartel,D.P., and Kingston,R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Leatherman,J.L. and Jongens,T.A. (2003). Transcriptional silencing and translational control: key features of early germline development. Bioessays *25*, 326-335.

Li,C., Vagin,V.V., Lee,S., Xu,J., Ma,S., Xi,H., Seitz,H., Horwich,M.D., Syrzycka,M., Honda,B.M. *et al.* (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509-521.

Li,E., Beard,C., and Jaenisch,R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

Li,E., Bestor,T.H., and Jaenisch,R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li,J., Wang,J., Wang,J., Nawaz,Z., Liu,J.M., Qin,J., and Wong,J. (2000). Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. EMBO J. *19*, 4342-4350.

Li,J.Y., English,M.A., Ball,H.J., Yeyati,P.L., Waxman,S., and Licht,J.D. (1997). Sequence-specific DNA binding and transcriptional regulation by the promyelocytic leukemia zinc finger protein. J. Biol. Chem. *272*, 22447-22455.

Lin,H. and Spradling,A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.

Lin,R.J., Nagy,L., Inoue,S., Shao,W., Miller,W.H., Jr., and Evans,R.M. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature *391*, 811-814.

Lin,W., Lai,C.H., Tang,C.J., Huang,C.J., and Tang,T.K. (1999). Identification and gene structure of a novel human PLZF-related transcription factor gene, TZFP. Biochem. Biophys. Res. Commun. *264*, 789-795.

Lister,R., Pelizzola,M., Dowen,R.H., Hawkins,R.D., Hon,G., Tonti-Filippini,J., Nery,J.R., Lee,L., Ye,Z., Ngo,Q.M. *et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.

Liu,D., Matzuk,M.M., Sung,W.K., Guo,Q., Wang,P., and Wolgemuth,D.J. (1998). Cyclin A1 is required for meiosis in the male mouse. Nat. Genet. *20*, 377-380.

Liu,Y., Taverna,S.D., Muratore,T.L., Shabanowitz,J., Hunt,D.F., and Allis,C.D. (2007). RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena. Genes Dev. *21*, 1530-1545.

Loenarz,C. and Schofield,C.J. (2008). Expanding chemical biology of 2-oxoglutarate oxygenases. Nat. Chem. Biol. *4*, 152-156.

Luco,R.F., Pan,Q., Tominaga,K., Blencowe,B.J., Pereira-Smith,O.M., and Misteli,T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996-1000.

Lyon,M.F. (2003). Transmission ratio distortion in mice. Annu. Rev. Genet. *37*, 393-408.

Ma,L., Buchold,G.M., Greenbaum,M.P., Roy,A., Burns,K.H., Zhu,H., Han,D.Y., Harris,R.A., Coarfa,C., Gunaratne,P.H. *et al.* (2009). GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. PLoS. Genet. *5*, e1000635.

Maclean,J.A. and Wilkinson,M.F. (2005). Gene regulation in spermatogenesis. Curr. Top. Dev. Biol. *71*, 131-197.

Maier,S. and Olek,A. (2002). Diabetes: a candidate disease for efficient DNA methylation profiling. J. Nutr. *132*, 2440S-2443S.

Mangelsdorf,D.J., Thummel,C., Beato,M., Herrlich,P., Schutz,G., Umesono,K., Blumberg,B., Kastner,P., Mark,M., Chambon,P. *et al.* (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835-839.

Marcon,E., Babak,T., Chua,G., Hughes,T., and Moens,P.B. (2008). miRNA and piRNA localization in the male mammalian meiotic nucleus. Chromosome. Res. *16*, 243-260.

Marsit,C.J., McClean,M.D., Furniss,C.S., and Kelsey,K.T. (2006). Epigenetic inactivation of the SFRP genes is associated with drinking, smoking and HPV in head and neck squamous cell carcinoma. Int. J. Cancer *119*, 1761-1766.

Marston,A.L. and Amon,A. (2004). Meiosis: cell-cycle controls shuffle and deal. Nat. Rev. Mol. Cell Biol. *5*, 983-997.

Mattick,J.S., Taft,R.J., and Faulkner,G.J. (2010). A global view of genomic information--moving beyond the gene and the master regulator. Trends Genet. *26*, 21-28.

McKim,K.S. and Hayashi-Hagihara,A. (1998). mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev. *12*, 2932-2942.

Meikar,O., Da,R.M., Korhonen,H., and Kotaja,N. (2011). Chromatoid body and small RNAs in male germ cells. Reproduction. *142*, 195-209.

Meikar,O., Da,R.M., Liljenback,H., Toppari,J., and Kotaja,N. (2010). Accumulation of piRNAs in the chromatoid bodies purified by a novel isolation protocol. Exp. Cell Res. *316*, 1567-1575.

Melnick,A., Ahmad,K.F., Arai,S., Polinger,A., Ball,H., Borden,K.L., Carlile,G.W., Prive,G.G., and Licht,J.D. (2000). In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. Mol. Cell Biol. *20*, 6550-6567.

Miaw,S.C., Choi,A., Yu,E., Kishikawa,H., and Ho,I.C. (2000). ROG, repressor of GATA, regulates the expression of cytokine genes. Immunity. *12*, 323-333.

Miki,Y., Nishisho,I., Horii,A., Miyoshi,Y., Utsunomiya,J., Kinzler,K.W., Vogelstein,B., and Nakamura,Y. (1992). Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res. *52*, 643-645.

Mittal,A., Piyathilake,C., Hara,Y., and Katiyar,S.K. (2003). Exceptionally high protection of photocarcinogenesis by topical application of (--)-epigallocatechin-3-gallate in hydrophilic cream in SKH-1 hairless mouse model: relationship to inhibition of UVB-induced global DNA hypomethylation. Neoplasia. *5*, 555-565.

Mochizuki,K. and Gorovsky,M.A. (2004). Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. Genes Dev. *18*, 2068-2073.

Moldovan,G.L. and D'Andrea,A.D. (2009). How the fanconi anemia pathway guards the genome. Annu. Rev. Genet. *43*, 223-249.

Monesi,V. (1964). Ribonucleic acid synthesis during mitosis and meiosis in the mouse testis. J. Cell Biol. *22*, 521-532.

Monesi,V. (1965). Synthetic activities during spermatogenesis in the mouse RNA and protein. Exp. Cell Res. *39*, 197-224.

Moore,G.P. (1971). DNA-dependent RNA synthesis in fixed cells during spermatogenesis in mouse. Exp. Cell Res. *68*, 462-465.

Morgan,H.D., Santos,F., Green,K., Dean,W., and Reik,W. (2005). Epigenetic reprogramming in mammals. Hum. Mol. Genet. *14 Spec No 1*, R47-R58.

Morse,B., Rotherg,P.G., South,V.J., Spandorfer,J.M., and Astrin,S.M. (1988). Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. Nature *333*, 87-90.

Mruk,D.D. and Cheng,C.Y. (2004). Sertoli-Sertoli and Sertoli-germ cell interactions and their significance in germ cell movement in the seminiferous epithelium during spermatogenesis. Endocr. Rev. *25*, 747-806.

Muller,T.A., Meek,K., and Hausinger,R.P. (2010). Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites. DNA Repair (Amst) *9*, 58-65.

Muto,A., Hoshino,H., Madisen,L., Yanai,N., Obinata,M., Karasuyama,H., Hayashi,N., Nakauchi,H., Yamamoto,M., Groudine,M. *et al.* (1998). Identification of Bach2 as a B-cell-specific partner for small maf proteins that negatively regulate the immunoglobulin heavy chain gene 3' enhancer. EMBO J. *17*, 5734-5743.

Nadler,J.J. and Braun,R.E. (2000). Fanconi anemia complementation group C is required for proliferation of murine primordial germ cells. Genesis. *27*, 117-123.

Nathan,D., Ingvarsdottir,K., Sterner,D.E., Bylebyl,G.R., Dokmanovic,M., Dorsey,J.A., Whelan,K.A., Krsmanovic,M., Lane,W.S., Meluh,P.B. *et al.* (2006). Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev. *20*, 966-976.

Nelson,C.J., Santos-Rosa,H., and Kouzarides,T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell *126*, 905-916.

Nicklas,R.B. (1974). Chromosome segregation mechanisms. Genetics *78*, 205-213.

Niedzwiedz,W., Mosedale,G., Johnson,M., Ong,C.Y., Pace,P., and Patel,K.J. (2004). The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. Mol. Cell *15*, 607-620.

Nigg,E.A. (2001). Mitotic kinases as regulators of cell division and its checkpoints. Nat. Rev. Mol. Cell Biol. *2*, 21-32.

Nowak,S.J. and Corces,V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet. *20*, 214-220.

Oswald,J., Engemann,S., Lane,N., Mayer,W., Olek,A., Fundele,R., Dean,W., Reik,W., and Walter,J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475-478.

Ozawa,Y., Towatari,M., Tsuzuki,S., Hayakawa,F., Maeda,T., Miyata,Y., Tanimoto,M., and Saito,H. (2001). Histone deacetylase 3 associates with and represses the transcription factor GATA-2. Blood *98*, 2116-2123.

Ozer,A. and Bruick,R.K. (2007). Non-heme dioxygenases: cellular sensors and regulators jelly rolled into one? Nat. Chem. Biol. *3*, 144-153.

Pal-Bhadra,M., Leibovitch,B.A., Gandhi,S.G., Rao,M., Bhadra,U., Birchler,J.A., and Elgin,S.C. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.

Pan,Z., Sikandar,S., Witherspoon,M., Dizon,D., Nguyen,T., Benirschke,K., Wiley,C., Vrana,P., and Lipkin,S.M. (2008). Impaired placental trophoblast lineage differentiation in Alkbh1(-/-) mice. Dev. Dyn. *237*, 316-327.

Panning,B. and Jaenisch,R. (1998). RNA and the epigenetic regulation of X chromosome inactivation. Cell *93*, 305-308.

Paralkar,V.M., Weeks,B.S., Yu,Y.M., Kleinman,H.K., and Reddi,A.H. (1992). Recombinant human bone morphogenetic protein 2B stimulates PC12 cell differentiation: potentiation and binding to type IV collagen. J. Cell Biol. *119*, 1721-1728.

Parvinen,M. (2005). The chromatoid body in spermatogenesis. Int. J. Androl *28*, 189-201.

Peters,A.H., O'Carroll,D., Scherthan,H., Mechtler,K., Sauer,S., Schofer,C., Weipoltshammer,K., Pagani,M., Lachner,M., Kohlmaier,A. *et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-337.

Peters,L. and Meister,G. (2007). Argonaute proteins: mediators of RNA silencing. Mol. Cell *26*, 611-623.

Piazza,F., Costoya,J.A., Merghoub,T., Hobbs,R.M., and Pandolfi,P.P. (2004). Disruption of PLZP in mice leads to increased T-lymphocyte proliferation, cytokine production, and altered hematopoietic stem cell homeostasis. Molecular and Cellular Biology *24*, 10456-10469.

Pillai,R.S. and Chuma,S. (2012). piRNAs and their involvement in male germline development in mice. Dev. Growth Differ.

Pokholok,D.K., Harbison,C.T., Levine,S., Cole,M., Hannett,N.M., Lee,T.I., Bell,G.W., Walker,K., Rolfe,P.A., Herbolsheimer,E. *et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Print,C.G. and Loveland,K.L. (2000). Germ cell suicide: new insights into apoptosis during spermatogenesis. Bioessays *22*, 423-430.

Prokhortchouk,E. and Hendrich,B. (2002). Methyl-CpG binding proteins and cancer: are MeCpGs more important than MBDs? Oncogene *21*, 5394-5399.

Qamar,I., Park,E., Gong,E.Y., Lee,H.J., and Lee,K. (2009). ARR19 (androgen receptor corepressor of 19 kDa), an antisteroidogenic factor, is regulated by GATA-1 in testicular Leydig cells. J. Biol. Chem. *284*, 18021-18032.

Ragan,C., Mowry,B.J., and Bauer,D.C. (2012). Hybridization-based reconstruction of small non-coding RNA transcripts from deep sequencing data. Nucleic Acids Res.

Reuter,M., Berninger,P., Chuma,S., Shah,H., Hosokawa,M., Funaya,C., Antony,C., Sachidanandam,R., and Pillai,R.S. (2011). Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature *480*, 264-267.

Reuter,M., Chuma,S., Tanaka,T., Franz,T., Stark,A., and Pillai,R.S. (2009). Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat. Struct. Mol. Biol. *16*, 639-646.

Rhee,I., Bachman,K.E., Park,B.H., Jair,K.W., Yen,R.W., Schuebel,K.E., Cui,H., Feinberg,A.P., Lengauer,C., Kinzler,K.W. *et al.* (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature *416*, 552-556.

Richardson,B. (2003a). DNA methylation and autoimmune disease. Clin. Immunol. *109*, 72-79.

Richardson,B. (2003b). Impact of aging on DNA methylation. Ageing Res. Rev. *2*, 245-261.

Ringrose,L. and Paro,R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu. Rev. Genet. *38*, 413-443.

Ringvoll,J., Moen,M.N., Nordstrand,L.M., Meira,L.B., Pang,B., Bekkelund,A., Dedon,P.C., Bjelland,S., Samson,L.D., Falnes,P.O. *et al.* (2008). AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. Cancer Res. *68*, 4142-4149.

Ringvoll,J., Nordstrand,L.M., Vagbo,C.B., Talstad,V., Reite,K., Aas,P.A., Lauritzen,K.H., Liabakk,N.B., Bjork,A., Doughty,R.W. *et al.* (2006). Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. EMBO J. *25*, 2189-2198.

Ripamonti,U. and Reddi,A.H. (1992). Growth and morphogenetic factors in bone induction: role of osteogenin and related bone morphogenetic proteins in craniofacial and periodontal bone repair. Crit Rev. Oral Biol. Med. *3*, 1-14.

Ripamonti,U. and Reddi,A.H. (1997). Tissue engineering, morphogenesis, and regeneration of the periodontal tissues by bone morphogenetic proteins. Crit Rev. Oral Biol. Med. *8*, 154-163.

Rizzoti,K. and Lovell-Badge,R. (2007). SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. Development *134*, 3437-3448.

Roeder,G.S. (1997). Meiotic chromosomes: it takes two to tango. Genes Dev. *11*, 2600-2621.

Roman-Gomez,J., Jimenez-Velasco,A., Agirre,X., Cervantes,F., Sanchez,J., Garate,L., Barrios,M., Castillejo,J.A., Navarro,G., Colomer,D., Prosper,F., Heiniger,A., and Torres,A. (2005). Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia. Oncogene *24*, 7213-7223.

Roth,S.Y. and Allis,C.D. (1996). Histone acetylation and chromatin assembly: a single escort, multiple dances? Cell *87*, 5-8.

Rouget,C., Papin,C., Boureux,A., Meunier,A.C., Franco,B., Robine,N., Lai,E.C., Pelisson,A., and Simonelig,M. (2010). Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature *467*, 1128-1132.

Rubin,G.M., Kidwell,M.G., and Bingham,P.M. (1982). The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell *29*, 987-994.

Ruggiu,M., Speed,R., Taggart,M., McKay,S.J., Kilanowski,F., Saunders,P., Dorin,J., and Cooke,H.J. (1997). The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. Nature *389*, 73-77.

Russell,L.D. and Clermont,Y. (1977). Degeneration of germ cells in normal, hypophysectomized and hormone treated hypophysectomized rats. Anat. Rec. *187*, 347-366.

Saito,K., Nishida,K.M., Mori,T., Kawamura,Y., Miyoshi,K., Nagami,T., Siomi,H., and Siomi,M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev. *20*, 2214-2222.

Sassone-Corsi,P. (2002). Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science *296*, 2176-2178.

Schofield,C.J. and McDonough,M.A. (2007). Structural and mechanistic studies on the peroxisomal oxygenase phytanoyl-CoA 2-hydroxylase (PhyH). Biochem. Soc. Trans. *35*, 870-875.

Sedgwick,B., Bates,P.A., Paik,J., Jacobs,S.C., and Lindahl,T. (2007). Repair of alkylated DNA: recent advances. DNA Repair (Amst) *6*, 429-442.

Shaffer,A.L., Yu,X., He,Y., Boldrick,J., Chan,E.P., and Staudt,L.M. (2000). BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. Immunity. *13*, 199-212.

Shah,N.M., Groves,A.K., and Anderson,D.J. (1996). Alternative neural crest cell fates are instructively promoted by TGFbeta superfamily members. Cell *85*, 331-343.

Sheardown,S.A., Duthie,S.M., Johnston,C.M., Newall,A.E., Formstone,E.J., Arkell,R.M., Nesterova,T.B., Alghisi,G.C., Rastan,S., and Brockdorff,N. (1997). Stabilization of Xist RNA mediates initiation of X chromosome inactivation. Cell *91*, 99-107.

Shilatifard,A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu. Rev. Biochem. *75*, 243-269.

Shogren-Knaak,M., Ishii,H., Sun,J.M., Pazin,M.J., Davie,J.R., and Peterson,C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Shoji,M., Tanaka,T., Hosokawa,M., Reuter,M., Stark,A., Kato,Y., Kondoh,G., Okawa,K., Chujo,T., Suzuki,T. *et al.* (2009). The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Dev. Cell *17*, 775-787.

Shuster,E.O. and Byers,B. (1989). Pachytene arrest and other meiotic effects of the start mutations in Saccharomyces cerevisiae. Genetics *123*, 29-43.

Siomi,M.C., Sato,K., Pezic,D., and Aravin,A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. *12*, 246-258.

Solari,A.J. (1974). The behavior of the XY pair in mammals. Int. Rev. Cytol. *38*, 273-317.

Songe-Moller,L., van den Born,E., Leihne,V., Vagbo,C.B., Kristoffersen,T., Krokan,H.E., Kirpekar,F., Falnes,P.O., and Klungland,A. (2010). Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. Mol. Cell Biol. *30*, 1814-1827.

Soper,S.F., van der Heijden,G.W., Hardiman,T.C., Goodheart,M., Martin,S.L., de,B.P., and Bortvin,A. (2008). Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. Dev. Cell *15*, 285-297.

Sterner,D.E. and Berger,S.L. (2000). Acetylation of histones and transcription-related factors. Microbiol. Mol. Biol. Rev. *64*, 435-459.

Straussman,R., Nejman,D., Roberts,D., Steinfeld,I., Blum,B., Benvenisty,N., Simon,I., Yakhini,Z., and Cedar,H. (2009). Developmental programming of CpG island methylation profiles in the human genome. Nat. Struct. Mol. Biol. *16*, 564-571.

Szutorisz,H., Canzonetta,C., Georgiou,A., Chow,C.M., Tora,L., and Dillon,N. (2005). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. Mol. Cell Biol. *25*, 1804-1820.

Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L., and Rao,A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takahara,T., Ohsumi,T., Kuromitsu,J., Shibata,K., Sasaki,N., Okazaki,Y., Shibata,H., Sato,S., Yoshiki,A., Kusakabe,M. *et al.* (1996). Dysfunction of the Orleans reeler gene arising from exon skipping due to transposition of a full-length copy of an active L1 sequence into the skipped exon. Hum. Mol. Genet. *5*, 989-993.

Takai,D., Yagi,Y., Habib,N., Sugimura,T., and Ushijima,T. (2000). Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. Jpn. J. Clin. Oncol. *30*, 306-309.

Tan,K.A., De,G.K., Atanassova,N., Walker,M., Sharpe,R.M., Saunders,P.T., Denolet,E., and Verhoeven,G. (2005). The role of androgens in sertoli cell proliferation and functional maturation: studies in mice with total or Sertoli cell-selective ablation of the androgen receptor. Endocrinology *146*, 2674-2683.

Tanaka,S.S., Toyooka,Y., Akasu,R., Katoh-Fukui,Y., Nakahara,Y., Suzuki,R., Yokoyama,M., and Noce,T. (2000). The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes Dev. *14*, 841-853.

Tang,C.J., Chuang,C.K., Hu,H.M., and Tang,T.K. (2001). The zinc finger domain of Tzfp binds to the tbs motif located at the upstream flanking region of the Aie1 (aurora-C) kinase gene. J. Biol. Chem. *276*, 19631-19639.

Tang,C.J., Lin,C.Y., and Tang,T.K. (2006). Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. Dev. Biol. *290*, 398-410.

Taylor,D.R. and Ingvarsson,P.K. (2003). Common features of segregation distortion in plants and animals. Genetica *117*, 27-35.

Toscani,A., Mettus,R.V., Coupland,R., Simpkins,H., Litvin,J., Orth,J., Hatton,K.S., and Reddy,E.P. (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-myb. Nature *386*, 713-717.

Trewick,S.C., Henshaw,T.F., Hausinger,R.P., Lindahl,T., and Sedgwick,B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature *419*, 174-178.

Tsai,M.Y., Yeh,S.D., Wang,R.S., Yeh,S., Zhang,C., Lin,H.Y., Tzeng,C.R., and Chang,C. (2006). Differential effects of spermatogenesis and fertility in mice lacking androgen receptor in individual testis cells. Proc. Natl. Acad. Sci. U. S. A *103*, 18975-18980.

Tseng,T.C., Chen,S.H., Hsu,Y.P., and Tang,T.K. (1998). Protein kinase profile of sperm and eggs: cloning and characterization of two novel testis-specific protein kinases (AIE1, AIE2) related to yeast and fly chromosome segregation regulators. DNA Cell Biol. *17*, 823-833.

Tsujikawa,K., Koike,K., Kitae,K., Shinkawa,A., Arima,H., Suzuki,T., Tsuchiya,M., Makino,Y., Furukawa,T., Konishi,N., and Yamamoto,H. (2007). Expression and sub-cellular localization of human ABH family molecules. J. Cell Mol. Med. *11*, 1105-1116.

Tsukada,Y., Fang,J., Erdjument-Bromage,H., Warren,M.E., Borchers,C.H., Tempst,P., and Zhang,Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tsuzuki,S. and Enver,T. (2002). Interactions of GATA-2 with the promyelocytic leukemia zinc finger (PLZF) protein, its homologue FAZF, and the t(11;17)-generated PLZF-retinoic acid receptor alpha oncoprotein. Blood *99*, 3404-3410.

Turner,B.M. (2007). Defining an epigenetic code. Nat. Cell Biol. *9*, 2-6.

Turner,J.M., Mahadevaiah,S.K., Ellis,P.J., Mitchell,M.J., and Burgoyne,P.S. (2006). Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. Dev. Cell *10*, 521-529.

Turner,J.M., Mahadevaiah,S.K., Fernandez-Capetillo,O., Nussenzweig,A., Xu,X., Deng,C.X., and Burgoyne,P.S. (2005). Silencing of unsynapsed meiotic chromosomes in the mouse. Nat. Genet. *37*, 41-47.

Unhavaithaya,Y., Hao,Y., Beyret,E., Yin,H., Kuramochi-Miyagawa,S., Nakano,T., and Lin,H. (2009). MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. J. Biol. Chem. *284*, 6507-6519.

Vagin,V.V., Klenov,M.S., Kalmykova,A.I., Stolyarenko,A.D., Kotelnikov,R.N., and Gvozdev,V.A. (2004). The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of Drosophila melanogaster. RNA. Biol. *1*, 54-58.

Vainio,S., Karavanova,I., Jowett,A., and Thesleff,I. (1993). Identification of BMP-4 as a signal mediating secondary induction between epithelial and mesenchymal tissues during early tooth development. Cell *75*, 45-58.

van den Born,E., Vagbo,C.B., Songe-Moller,L., Leihne,V., Lien,G.F., Leszczynska,G., Malkiewicz,A., Krokan,H.E., Kirpekar,F., Klungland,A. *et al.* (2011). ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA. Nat. Commun. *2*, 172.

Verhoeven,G. (1992). Local control systems within the testis. Baillieres Clin. Endocrinol. Metab *6*, 313-333.

Verhoeven,G., Willems,A., Denolet,E., Swinnen,J.V., and De,G.K. (2010). Androgens and spermatogenesis: lessons from transgenic mouse models. Philos. Trans. R. Soc. Lond B Biol. Sci. *365*, 1537-1556.

Viger,R.S., Taniguchi,H., Robert,N.M., and Tremblay,J.J. (2004). Role of the GATA family of transcription factors in andrology. J. Androl *25*, 441-452.

Vourekas,A., Zheng,Q., Alexiou,P., Maragkakis,M., Kirino,Y., Gregory,B.D., and Mourelatos,Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat. Struct. Mol. Biol. *19*, 773-781.

Walsh,C.P., Chaillet,J.R., and Bestor,T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. *20*, 116-117.

Ward,W.S. and Coffey,D.S. (1991). DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. Biol. Reprod. *44*, 569-574.

Watamoto,K., Towatari,M., Ozawa,Y., Miyata,Y., Okamoto,M., Abe,A., Naoe,T., and Saito,H. (2003). Altered interaction of HDAC5 with GATA-1 during MEL cell differentiation. Oncogene *22*, 9176-9184.

Watanabe,T., Totoki,Y., Toyoda,A., Kaneda,M., Kuramochi-Miyagawa,S., Obata,Y., Chiba,H., Kohara,Y., Kono,T., Nakano,T. *et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Watson,E.D., Mattar,P., Schuurmans,C., and Cross,J.C. (2009). Neural stem cell self-renewal requires the Mrj co-chaperone. Dev. Dyn. *238*, 2564-2574.

Wei,Y.F., Carter,K.C., Wang,R.P., and Shell,B.K. (1996). Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. Nucleic Acids Res. *24*, 931-937.

Westbye,M.P., Feyzi,E., Aas,P.A., Vagbo,C.B., Talstad,V.A., Kavli,B., Hagen,L., Sundheim,O., Akbari,M., Liabakk,N.B. *et al.* (2008). Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J. Biol. Chem. *283*, 25046-25056.

Wolffe,A.P. and Pruss,D. (1996). Targeting chromatin disruption: Transcription regulators that acetylate histones. Cell *84*, 817-819.

Wong,C.W. and Privalsky,M.L. (1998). Transcriptional repression by the SMRT-mSin3 corepressor: multiple interactions, multiple mechanisms, and a potential role for TFIIB. Mol. Cell Biol. *18*, 5500-5510.

Wong,J.C., Alon,N., Mckerlie,C., Huang,J.R., Meyn,M.S., and Buchwald,M. (2003). Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. Hum. Mol. Genet. *12*, 2063-2076.

Wordinger,R.J. and Clark,A.F. (2007). Bone morphogenetic proteins and their receptors in the eye. Exp. Biol. Med. (Maywood.) *232*, 979-992.

Wyatt,A.W., Osborne,R.J., Stewart,H., and Ragge,N.K. (2010). Bone morphogenetic protein 7 (BMP7) mutations are associated with variable ocular, brain, ear, palate, and skeletal anomalies. Hum. Mutat. *31*, 781-787.

Xiong,W.C. and Montell,C. (1993). tramtrack is a transcriptional repressor required for cell fate determination in the Drosophila eye. Genes Dev. *7*, 1085-1096.

Xu,L., Wei,Y., Reboul,J., Vaglio,P., Shin,T.H., Vidal,M., Elledge,S.J., and Harper,J.W. (2003). BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. Nature *425*, 316-321.

Yang,K.T., Li,S.K., Chang,C.C., Tang,C.J., Lin,Y.N., Lee,S.C., and Tang,T.K. (2010). Aurora-C kinase deficiency causes cytokinesis failure in meiosis I and production of large polyploid oocytes in mice. Mol. Biol. Cell *21*, 2371-2383.

Yang,X.J. and Seto,E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene *26*, 5310-5318.

Yang,Y., Tse,A.K., Li,P., Ma,Q., Xiang,S., Nicosia,S.V., Seto,E., Zhang,X., and Bai,W. (2011). Inhibition of androgen receptor activity by histone deacetylase 4 through receptor SUMOylation. Oncogene *30*, 2207-2218.

Yomogida,K., Ohtani,H., Harigae,H., Ito,E., Nishimune,Y., Engel,J.D., and Yamamoto,M. (1994). Developmental stage- and spermatogenic cycle-specific expression of transcription factor GATA-1 in mouse Sertoli cells. Development *120*, 1759-1766.

Yoon,H.S., Scharer,C.D., Majumder,P., Davis,C.W., Butler,R., Zinzow-Kramer,W., Skountzou,I., Koutsonanos,D.G., Ahmed,R., and Boss,J.M. (2012). ZBTB32 Is an Early Repressor of the CIITA and MHC Class II Gene Expression during B Cell Differentiation to Plasma Cells. J. Immunol.

Zardo,G., Ciolfi,A., Vian,L., Starnes,L.M., Billi,M., Racanicchi,S., Maresca,C., Fazi,F., Travaglini,L., Noguera,N. *et al.* (2012). Polycombs and microRNA-223 regulate human granulopoiesis by transcriptional control of target gene expression. Blood *119*, 4034-4046.

Zhang,Y. and Reinberg,D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343-2360.

Zhao,G.Q., Deng,K., Labosky,P.A., Liaw,L., and Hogan,B.L. (1996). The gene encoding bone morphogenetic protein 8B is required for the initiation and maintenance of spermatogenesis in the mouse. Genes Dev. *10*, 1657-1669.

Zhao,G.Q. and Hogan,B.L. (1996). Evidence that mouse Bmp8a (Op2) and Bmp8b are duplicated genes that play a role in spermatogenesis and placental development. Mech. Dev. *57*, 159-168.

Zickler,D. and Kleckner,N. (1999). Meiotic chromosomes: integrating structure and function. Annu. Rev. Genet. *33*, 603-754.

Ziegelbauer,J., Shan,B., Yager,D., Larabell,C., Hoffmann,B., and Tjian,R. (2001). Transcription factor MIZ-1 is regulated via microtubule association. Mol. Cell *8*, 339-349.

Zou,H. and Niswander,L. (1996). Requirement for BMP signaling in interdigital apoptosis and scale formation. Science *272*, 738-741.

Appendices; Papers and Manuscripts

I Nordstrand L.M., Svärd J.*, Larsen E.*, Nilsen A.*, Ougland R.*, **Furu K.***, Lien GF, Rognes T, Namekawa SH, Lee JT, Klungland A. "Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects" (* contributed equally)

PLoS One. 2010 Nov 3;5(11):e13827

II Nordstrand L.M.*, **Furu K.***, Paulsen J., Rognes T., Klungland A. "Alkbh1 and Tzfp Repress a Non-Repeat piRNA Cluster in Pachytene Spermatocytes" (* contributed equally)

Nucleic Acids Research. 2012, in press

III **Furu K.**, Klungland A. "Tzfp represses the Androgen Receptor in mouse testis"

Manuscript (2012)

I

OPEN ACCESS Freely available online

PLoS one

Mice Lacking Alkbh1 Display Sex-Ratio Distortion and Unilateral Eye Defects

Line M. Nordstrand[1], Jessica Svärd[1☯], Elisabeth Larsen[1,2☯], Anja Nilsen[1☯], Rune Ougland[1,2☯], Kari Furu[1☯], Guro F. Lien[1], Torbjørn Rognes[1,3], Satoshi H. Namekawa[2], Jeannie T. Lee[2,5], Arne Klungland[1,4]*

1 Centre for Molecular Biology and Neuroscience, Institute of Medical Microbiology, Oslo University Hospital and University of Oslo, Oslo, Norway, **2** Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **3** Department of Informatics, University of Oslo, Oslo, Norway, **4** Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway, **5** Howard Hughes Medical Institute, Massachusetts General Hospital, Boston, Massachusetts, United States of America

Abstract

Background: *Eschericia coli* AlkB is a 2-oxoglutarate- and iron-dependent dioxygenase that reverses alkylated DNA damage by oxidative demethylation. Mouse AlkB homolog 1 (Alkbh1) is one of eight members of the newly discovered family of mammalian dioxygenases.

Methods and Findings: In the present study we show non-Mendelian inheritance of the *Alkbh1* targeted allele in mice. Both *Alkbh1*$^{-/-}$ and heterozygous *Alkbh1*$^{+/-}$ offspring are born at a greatly reduced frequency. Additionally, the sex-ratio is considerably skewed against female offspring, with one female born for every three to four males. Most mechanisms that cause segregation distortion, act in the male gametes and affect male fertility. The skewing of the sexes appears to be of paternal origin, and might be set in the pachythene stage of meiosis during spermatogenesis, in which *Alkbh1* is upregulated more than 10-fold. In testes, apoptotic spermatids were revealed in 5–10% of the tubules in *Alkbh1*$^{-/-}$ adults. The deficiency of *Alkbh1* also causes misexpression of *Bmp2*, *4* and *7* at E11.5 during embryonic development. This is consistent with the incompletely penetrant phenotypes observed, particularly recurrent unilateral eye defects and craniofacial malformations.

Conclusions: Genetic and phenotypic assessment suggests that Alkbh1 mediates gene regulation in spermatogenesis, and that Alkbh1 is essential for normal sex-ratio distribution and embryonic development in mice.

Citation: Nordstrand LM, Svärd J, Larsen E, Nilsen A, Ougland R, et al. (2010) Mice Lacking Alkbh1 Display Sex-Ratio Distortion and Unilateral Eye Defects. PLoS ONE 5(11): e13827. doi:10.1371/journal.pone.0013827

Editor: Joseph Najbauer, City of Hope National Medical Center, United States of America

Received May 10, 2010; **Accepted** October 14, 2010; **Published** November 3, 2010

Funding: This work was supported by the National Program in Functional Genomics (FUGE) sponsored by the Norwegian Research Council, the Norwegian Cancer Society and the European Union program 'DNA repair'. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: arne.klungland@rr-research.no

☯ These authors contributed equally to this work.

Introduction

The *Eschericia coli* (*E. coli*) DNA repair enzyme AlkB demethylates e.g. 1-methyladenine (1-meA) to adenine – generating succinate and formaldehyde – in the presence of iron as cofactor and 2-oxoglutarate as cosubstrate [1,2]. To date, eight AlkB homologs have been identified in the mammalian genome [3]. Except for Alkbh5, all the remaining proteins have been identified throughout the animal kingdom, suggesting fundamental roles in biological processes [4]. Two of these homologs, ALKBH2 and ALKBH3 in humans (Alkbh2 and Alkbh3 in mice), are similar to *E. coli* AlkB in that they efficiently repair damaged nucleic acids in the presence of iron and 2-oxoglutarate *in vitro* [5–9]. In mice, Alkbh2 is the major, probably only, dioxygenase that repairs 1-meA DNA *in vivo* and mice lacking Alkbh2 accumulate 1-meA in the genome during ageing [10]. This year, two groups reported that Alkbh8 is a tRNA methyltransferase required for the final step in the biogenesis of mcm^{5}U [11,12]. ALKBH8 plays important roles in the survival and progression of human bladder cancer both *in vitro* and *in vivo* [13]. A likely ninth AlkB homolog, the obesity-associated Fto protein, was shown to have potential to demethylate 3-methylthymine (3-meT) [14,15]. Crystal structure of the FTO protein recently confirmed this, and indicated that single-stranded RNA is the primary substrate of FTO [16]. Similarly, recombinant truncated Alkbh1 enzyme may demethylate 3-methylcytosine *in vitro* [17], but it remains unclear whether this activity is physiologically relevant.

All eight mammalian AlkB homologs contain the conserved iron- and 2-oxoglutarate dioxygenase domain. However, the region of *E. coli* AlkB that interacts with the nucleic acid substrate, the N-terminal nucleotide recognition lid, does not share sequence similarity with the mammalian homologs. Therefore, one cannot exclude the possibility that the targets of such proteins are not nucleic acids, but other macromolecules such as proteins. Since JmjC histone demethylases remove methyl groups from histones using the same mechanism as *E. coli* AlkB, it has been suggested that Alkbh1, 4 and 7 might be involved in histone/protein demethylation [18,19]. However, for Alkbh1 we, and others, have been unable to identify DNA/histone demethylation activity

[6,7,20,21]. In 2008 a paper on Alkbh1 was published by Pan et al, where a gene-targeting study in mice showed that Alkbh1 localizes to nuclear euchromatin and functions in epigenetic regulation of gene expression [20]. Their study demonstrated impaired placental trophoblast lineage differentiation in *Alkbh1*$^{-/-}$ mice, and a strong interaction of Alkbh1 with Mrj, an essential placental gene that mediates gene repression by recruitment of class II histone deacetylases (HDAC) [20].

In the present study we attempt to elucidate the role of Alkbh1 by targeted deletion in C57/BL6 mice. We demonstrate that *Alkbh1* deficiency in mice results in apoptosis in adult testes and sex-ratio distortion of offspring, most likely caused by defects in the pachytene stage during spermatogenesis. An incompletely penetrant phenotype apparent during embryonic development is consistent with *Bmp2*, *4* and *7* misexpression. Although many mechanistic aspects of Alkbh1 function remain to be revealed, these results show that Alkbh1 is crucial for normal embryonic development and viability in mice, and plays an important role during spermatogenesis.

Materials and Methods

Generation of *Alkbh1* Targeted Mice

A specific 360-bp murine probe of exon 6 in the *Alkbh1* gene was amplified from mouse genomic DNA by polymerase chain reaction (PCR) and used to screen a 129 SvJ mouse genomic library (Stratagene). To generate the targeting construct, we subcloned fragments from a ~14-kb genomic clone on both sides of neomycin *(neo)* in the pGT-N38 vector (New England Biolabs). Homologous arms consisting of a 3.0-kb MfeI/HindIII fragment and a 3.7-kb BsrGI fragment facilitated removal of a 3.8-kb HindIII/BsrGI fragment including exon 6 and replacement with the *neo* cassette. The targeting construct was electroporated into 129 SvJ embryonic stem (ES) cells, and transfectants were selected in geneticin (G418) and expanded for further analysis. Chimaeric mice were produced by microinjection of one targeted ES cell clone with normal karyotype into C57/BL6 blastocysts at embryonic day 3.5 (E3.5). We verified germline transmission of the targeted allele by Southern-blot analysis of ScaI-digested genomic DNA on the 5′ end and PCR analysis on the 3′ end. 5′ and 3′ homologous recombination in the F_1 generation were confirmed by PCR analysis. Heterozygous males were backcrossed for three generations onto C57/BL6 females. All mouse experiments were approved by the Norwegian Animal Research Authority (Ref. nr. 08/9940) and done in accordance with institutional guidelines at the Centre for Comparative Medicine at Oslo University Hospital. Animal work was conducted in accordance with the rules and regulations of the Federation of European Laboratory Animal Science Association's (FELASA).

Genotyping

For *Alkbh1* genotyping, ear-clip samples were degraded by incubation in PBND buffer (50 mM KCl, 10 mM Tris-HCl pH 8.3, 2.5 mM $MgCl_2$-$6H_2O$, 0.1 mg/ml gelatin, 0.45% v/v NP40, 0.45% v/v Tween 20) and 0.5 mg/ml proteinase K at 55°C over night. Samples were heated to 95°C for 10 min to inactivate proteinase K, and PCR amplified for 35 cycles with an annealing temperature of 60°C (see primers below). For sex genotyping of embryos, a small piece of tissue was obtained from the embryosac or -tail and washed three times in PBS to eliminate maternal contamination. The tissue was degraded by a 3-hour incubation, and subsequently treated as above. PCR analysis of *Sry* (Y-linked gene) was performed to determine maleness and *Rapsn* was used as an autosomal, internal control as described (Mouse Phenotypes, a Handbook of Mutatation Analysis, Cold Spring Harbor laboratory press, Chapter 3, page 40, 2005).

Primers wild-type allele (WT): 5′-AGTTATCAGGGCCATC-CAGGGAGGT-3′

5′-AACTGAGAGGTACAGGGAAGCATAA-3′

Primers targeted allele (KO): 5′-GCTTGCCGAATATCAT-GGTG-3′

5′-AACTGAGAGGTACAGGGAAGCATAA-3′

Whole-Mount *In Situ* Hybridization

We carried out whole-mount *in situ* hybridization on E9.5 to E12.5 embryos fixed in paraformaldehyde as described (Henrique et al. 1995). Mouse antisense and sense (control) RNA probes were prepared using DIG RNA labeling mix (Roche) together with T3 or Sp6 and T7 RNA polymerases (Roche). Templates for the labeling reaction were PCR products amplified from full-length mouse cDNA with T3, Sp6 or T7 promoters added to the PCR primers. For *Alkbh1* the template contained 465-bp of exon 6, for *Bmp2* 519-bp of exon 2–3 and for *Bmp7* 559-bp of exon 2–5. For *Bmp4*, linearized pSP72 plasmid with a 1550-bp insert was used as template. Embryos were examined on a SMZ1500 microscope (Nikon).

Quantitative Real-Time PCR (qPCR) Analysis

Total RNA was isolated from embryos, organs and germ cells using the Fast RNA Pro Green Kit (MP Biomedicals) according to the manufacturers protocol. Any DNA remnants were removed using TURBO DNase (Ambion) and cDNA was made using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The quantitative PCR reactions were carried out on a StepOnePlus or 7500 Fast instrument using 50 ng cDNA, TaqMan® Fast Universal PCR Master Mix and appropriate TaqMan primers and probes (all from Applied Biosystems). Pre-designed primers and probes were used both for the target genes (*Alkbh1, Vav2, Mapk8, Ccdc80, Rest, and Hif1a*) and endogenous controls (*Gapdh, 18s* and *β-actin*). All samples were run in triplicates and with one technical parallel (2 runs per sample). The relative quantity was calculated using the equation $RQ = 2^{-\Delta\Delta CT}$, where RQ is the relative quantity of the target gene. ΔΔCT is the difference in CT-value between the target gene and the endogenous control minus the difference in CT-values between the reference gene and the endogenous control.

STAPUT Isolation of Testicular Cells

Male germ cells were isolated from testes using an adapted version of the STAPUT method [22]. Pachytene cells and round spermatids were isolated from six 12-week old males, while a total of sixty 10-day old males were sacrificed for the isolation of type A and type B spermatogonia. The testes were put in ice cold DMEM medium containing antibiotics and then carefully detunicated. The tubules were treated with DNaseI, collagenase, trypsin and hyaluronidase (all from Sigma-Aldrich) at 34°C to remove connective tissue and somatic cells, yielding a cell suspension of germinal cells in DMEM containing 0.5% BSA. The cell suspension was loaded into the cell loading chamber of the STAPUT apparatus and separated by sedimentation velocity at unit gravity in a 2–4% w/v BSA gradient in DMEM medium at 4°C for 2.5 hours. After sedimentation, 10 ml fractions were collected and checked under the microscope. Fractions containing pure germ cells were pooled and the cell number counted in a *Countess*® Automated Cell Counter (Invitrogen). Cells were spun down and the pellet was snap frozen in liquid nitrogen before placed in −70°C. An aliquot of isolated cells was fixed on SuperFrost Plus slides (VWR) using Cell Adherence Solution (Crystalgen, Lot no 425081) for microscopic analysis of purity.

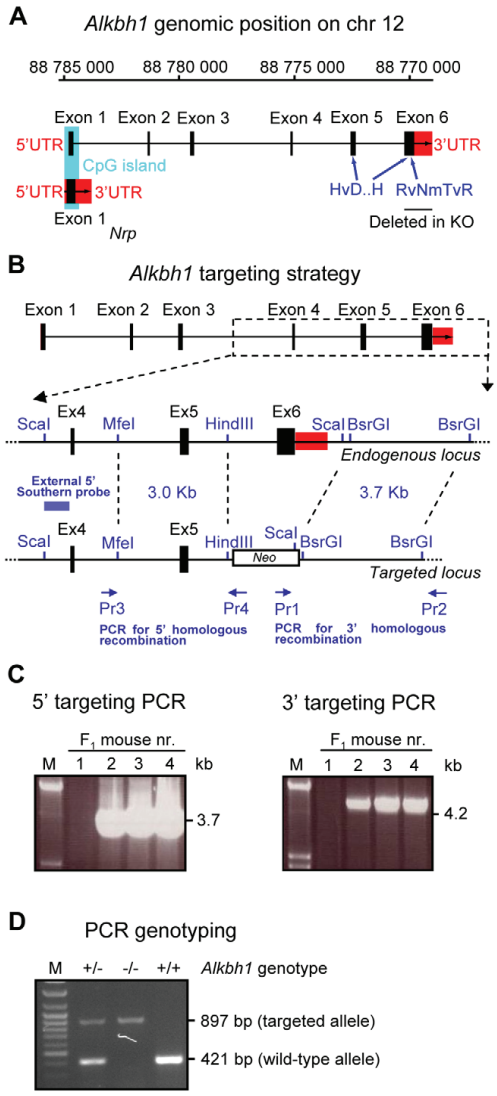


Figure 1. Targeted disruption of *Alkbh1* in embryonic stem cells and mice. (A) Schematic representation of the genomic region harboring the *Alkbh1* gene. Exon 6 is replaced by neomycin, thus maintaining the overlapping *Nrp* gene, and removing the conserved 2-oxoglutarate interaction domain (RvNmTvR) and parts of the iron-binding cluster (HvD…H) essential for enzymatic activity. A CpG island of 550 bp, shown in blue, is surrounding exon 1 (criteria used: Island size >200, GC Percent >50.0, Obs/Exp >0.6). The *Nrp* gene displays an overlap with exon 1 and is encoded as a forward frameshift to *Alkbh1* in the mouse. The 5′ and 3′ UTRs of *Alkbh1* and *Nrp* are shown in red. Coding exons are shown as black boxes. (B) Overview of the *Alkbh1* targeting strategy. Upper, schematic map of the genomic *Alkbh1* locus. Dashed lines point out the region used for homologous recombination. Middle, partial resctriction map of the endogenous *Alkbh1* locus participating in homologous recombination. Bottom, the targeted *Alkbh1* locus after correct integration of genomic fragments consisting of a 3.0-kb MfeI/HindIII fragment and a 3.7-kb BsrGI fragment on both sides of *Neo*, thereby replacing a 3.8-kb HindIII/BsrGI fragment including exon 6 with *Neo*. (C) PCR analysis for verification of 5′ and 3′ homologous recombination in the F_1 generation. The 3.7-kb 5′ targeted band (Pr3, Pr4) and the 4.2-kb 3′ targeted band (Pr1, Pr2) is present in F_1 mouse nr. 2, 3 and 4. M is the DNA marker. (D) PCR genotyping of the *Alkbh1* allele. The 421-bp wild-type band (WT) and the 897-bp targeted band (KO) is shown. M is the DNA marker.
doi:10.1371/journal.pone.0013827.g001

TUNEL Assay of Testes

We fixed testes from 3- and 9-month old animals in neutral-buffered formalin, progressively dehydrated them in a graded ethanol series, and embedded them in paraffin. Sections (4-µm) were deparaffinized and treated with proteinase K for 15 min and quenched in 3% hydrogen peroxide in PBS for 5 min at room temperature. Subsequently, nuclear staining in apoptotic cells was detected using ApopTag kit (Chemicon, http://www.chemicon.com) according to the manufacturers instruction. Sections were analysed on an Axioplan 2 microscope (Zeiss).

Immunofluorescent Staining of Testicular Cells

Testicular cells from 12-month old males were spread on SuperFrost Plus slides (VWR), progressively dehydrated in a graded ethanol series and dried completely. Slides were washed in 1× PBS and fixated in 4% PFA in PBS. Slides were blocked in 5% serum in PBS for 1 hour at room temperature and incubated with primary antibodies overnight at 4°C prior to detection with secondary antibodies. Primary antibodies used were rabbit anti-MacroH2A (1:500, Upstate) and mouse anti-FK2 (1:5000, Biomol). Secondary antibodies used were goat anti-rabbit Alexa Fluor 488 (green dye) (Invitrogen) and goat anti-mouse Alexa Fluor 594 (red dye) (Invitrogen), respectively. Single Z-sections were captured by Axioplan 2 microscope (Zeiss).

DNA Microarray Analysis

High quality of total RNA extracted from adult testes was verified on Agilent Bioanalyzer 2100 (RIN value between 9.8 and 10.0). 15 µg of biotinylated and fragmented cRNA was then hybridized onto the GeneChip Mouse Genome 430 2.0 Array (Affymetrix) according to manufacturers protocols (Affymetrix). QCs including scale factor, background, noise, spikes and RNA degradation were checked and validated using the yaqcaffy library (http://www.bioconductor.org/packages/2.3/%20bioc/vignettes/yaqcaffy/inst/doc/yaqcaffy.pdf).

Affymetrix raw data was generated with GCOS 1.4 (GeneChip Operating Software, Affymetrix), and the signal intensities of each probe set were normalized with the RMA (Robust Microarray Anlaysis) algorithm. To find differentially expressed genes, t-test with randomized variance was used as statistical test and the cut-off (p-value) was set to 0.05 with a FDR correction. Class comparison analysis was used to identify interesting genes. The signal in one group was always (i.e. for all the triplicate) higher or lower compared to the other group. Fold change for all the genes that passed the above criteria was computed and only the genes with ≥2-fold change were studied. The heatmap was generated using the GeneSpring GX 10 demoversion (Agilent). All data is MIAME compliant and the raw data has been deposited in a MIAME compliant database, the accession number is GSE22073.

Skeletal Staining

For skeletal analysis, skin and internal viscera of E18.5 embryos and newborn mice were removed. We then fixed the animals overnight in 95% ethanol and carried out Alcian blue 8GX

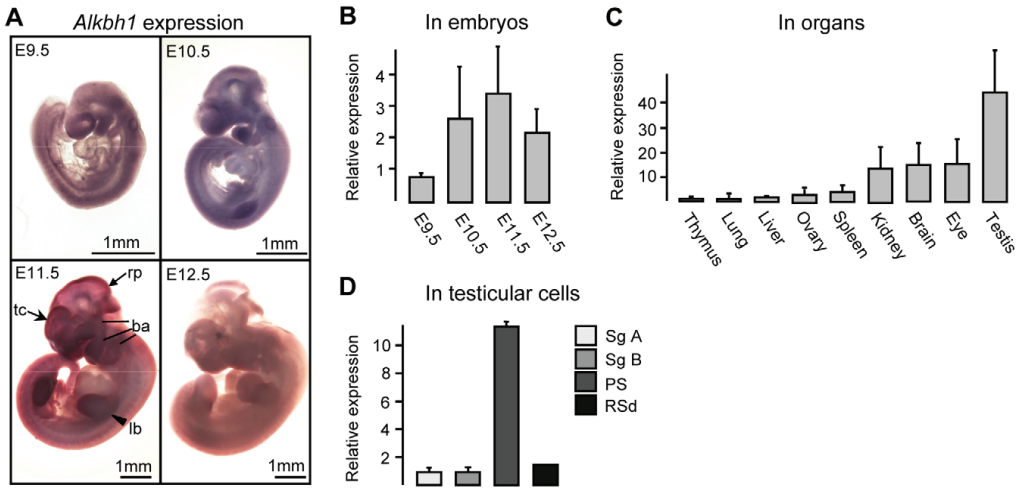


Figure 2. Expression of *Alkbh1* in embryos, organs and male germ cells. (A) Whole-mount *in situ* hybridization of *Alkbh1* between E9.5 and E12.5, side view. Peak expression is revealed at E11.5 in the telencephalon (tc) and frontonasal process, the maxillary and mandibular and hyoid arches (ba), the upper and lower limb buds (lb), and the midbrain and rhombomere 1 (r1) roof plates (rp). (B) Expression of *Alkbh1* between E9.5 and E12.5 by qPCR in RNA extracted from three - eight whole embryos. Peak expression at E11.5 is confirmed. Reference sample, E9.5 (RQ = 1.00); endogenous control, *Gapdh*. (C) Expression of *Alkbh1* in mouse organs by qPCR in RNA extracted from three - five 12-week old mice. Peak expression is shown in testis. Reference sample, thymus (RQ = 1.00); endogenous control, *18s*. (D) Expression of *Alkbh1* at different stages of spermatogenesis. Male germ cells from C57/BL6 mice were STAPUT sorted into type A spermatogonia (Sg A), type B spermatogonia (Sg B), pachytene spermatocytes (PS) and round spermatids (RSd), and analysed by qPCR after RNA extraction from the purified cell populations. Reference sample, type A spermatogonia (RQ = 1.00); endogenous control, *β-actin*.
doi:10.1371/journal.pone.0013827.g002

(Sigma) and Alizarin red S (Sigma) staining of cartilage and bone, respectively, as described (Manipulating the Mouse Embryo, 3rd edition, Cold Spring Harbor laboratory press, Chapter 16, Protocol 22, pages 699–700, 2003). The skeletons were photographed with a Nikon D80 camera.

Histological Analysis of Eyes

We fixed adult eyes in neutral-buffered formalin or paraformaldehyde added 20% absolute alcohol for 24 hours, progressively dehydrated them in a graded ethanol series, and embedded them in paraffin. Sections (4-μm) were deparaffinized, rehydrated and stained with hematoxylin and eosin. Sections were analysed on an AxioCam HRc microscope (Zeiss).

Results

Deletion of *Alkbh1* in Embryonic Stem Cells and Mice

To gain more insight into the role of the Alkbh1 dioxygenase we have generated mice lacking *Alkbh1*. Alkbh1 was the first mammalian AlkB homolog to be identified [23], and is the AlkB homolog most similar in sequence to *Eschericia coli (E. coli)* AlkB. The region of greatest similarity includes 107 amino acids, 37% of which are identical between the *E. coli* and mouse Alkbh1. The conserved RvNmTvR and HvD…H motifs of the 2-oxoglutarate and iron binding sites, respectively, are also present in both proteins. The conserved domains of Alkbh1 are encoded by exon 5 and 6 at the 3′ end of the mouse *Alkbh1* gene. To fully eliminate the activity of Alkbh1 and keep the overlapping *Nrp* gene intact, we substituted exon 6 with a neomycin-resistance gene cassette by homologous recombination in mouse embryonic stem cells (Figs. 1A–D). The expression of the *Nrp* gene was confirmed by qPCR (data not shown).

Expression Analysis of *Alkbh1* in Embryos, Organs and Male Germ Cells

The expression pattern of *Alkbh1* was analysed in embryos at different stages by whole-mount *in situ* hybridization (Fig. 2A) and by qPCR (Fig. 2B). Weak expression of *Alkbh1* was observed throughout the embryo at E8.5 (data not shown). As the cells migrate and differentiate during organogenesis the expression becomes more specific, and *Alkbh1* was detected in the spinal cord, forebrain and branchial arches at E9.5, and also in limb buds at E10.5 (Fig. 2A). Peak expression was detected at E11.5 in the frontonasal process including telencephalon (tc), maxillary, mandibular and hyoid arches (ba), upper and lower limb buds (lb), and midbrain and rhombomere 1 (r1) roof plates (rp) (Fig. 2A). *Alkbh1* expression decreased considerably from E11.5 to E12.5 (Fig. 2A–B). In adult organs, *Alkbh1* was highly expressed in testis (RQ = 44.0), with slightly lower expression in eye, brain and kidney (RQ = 16.0, 15.4, 14.4) (Fig. 2C). Moreover, the expression of *Alkbh1* was studied at different stages during spermatogenesis, and was found to be significantly elevated in the pachytene spermatocytes (PS) (RQ = 11.3) compared with spermatogonia A and B (Sg A, Sg B) and round spermatids (RSd) (RQ = 1.7) (Fig. 2D). This is the third stage of the prophase of meiosis I, in which synapsis is completed and homologous recombination occurs. Thus, Alkbh1 may have considerable potential for gene-function in embryonic development and in the pachytene stage during spermatogenesis.

Non-Mendelian Inheritance and Sex-Ratio Distortion in *Alkbh1* Targeted Mice

Mendelian inheritance, in which each parent contributes one of two possible alleles for a given trait, has a characteristic ratio of

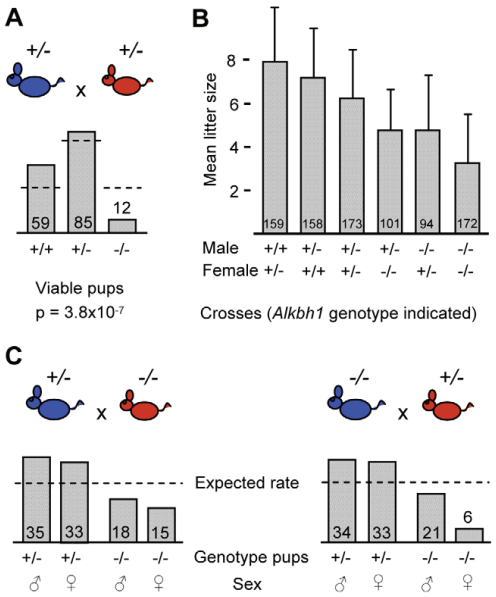


Figure 3. Non-Mendelian inheritance and sex-ratio distortion in *Alkbh1* targeted mice. (A) Offspring distribution of different genotypes at 1-month after crosses between heterozygous males (blue) and females (red) is shown (+ indicates wild-type *Alkbh1* allele and – indicates targeted *Alkbh1* allele). Mean number of pups per cross from 25 litters are represented. Dashed lines represent expected Mendelian distribution, and the χ^2-test was used to determine significance. (B) Average litter sizes from all crosses at 1-month of age are presented on the y-axis, while *Alkbh1* genotype of males and females used in the different crosses are indicated on the x-axis. The corresponding number of pups from more than 20 litters per cross are shown on the bars. (C) Left panel, crosses between heterozygous males (blue) and homozygous females (red). Right panel, crosses between homozygous males (blue) and heterozygous females (red). Offspring distribution of different *Alkbh1* genotypes at 1-month is shown. Mean number of pups per cross, calculated from 21 litters in the left panel and 20 litters in the right panel, are shown. Dashed lines represent expected sex-ratio distribution.
doi:10.1371/journal.pone.0013827.g003

1:2:1 after heterozygous crosses. Initial crosses of mice carrying either one or two targeted *Alkbh1* loci revealed non-Mendelian distribution. Therefore, we carried out extensive breeding analysis and genotyped more than 1400 *Alkbh1* mutant mice and embryos (Fig. 3). Following heterozygous breedings, the survival of $Alkbh1^{-/-}$ pups after 1 month was only 20% compared with wild-type littermates (Fig. 3A). In addition, the frequency of viable $Alkbh1^{+/-}$ mice was only 60% of the expected rate (Fig. 3A). The non-Mendelian distribution was clearly significant with a p-value of 3.8×10^{-7} (χ^2-test). A similar pattern was observed in $Alkbh1^{+/-}$ male x $Alkbh1^{-/-}$ female crosses, $p=5\times10^{-4}$ (χ^2-test) and $Alkbh1^{-/-}$ male x $Alkbh1^{+/-}$ female crosses, $p=3.7\times10^{-5}$ (χ^2-test) (Fig. 3C). In general, the average litter size decreased as the number of targeted alleles in the parental generation increased (Fig. 3B). The mean litter size was 9.2 for wild-type crosses, 6.2 for heterozygous crosses and 3.2 for homozygous crosses (Fig. 3B). Notably, paternal inheritance of the targeted allele seemed to be more critical than maternal transmission for the survival of offspring. Another evident phenotype was the growth retardation observed in viable $Alkbh1^{-/-}$ mice compared with wild-type littermates (Fig. S1).

One process of non-Mendelian inheritance is segregation distortion. There are a number of mechanisms that can cause segregation distortion, and both autosomal sex-ratio distortion as well as segregation distortion of the sex chromosomes exist [24]. In $Alkbh1^{+/-}$ male x $Alkbh1^{-/-}$ female crosses, the ratio of female to male homozygous offspring at 1 month was approximately 1:1 (Fig. 3C, left panel). In $Alkbh1^{-/-}$ male x $Alkbh1^{+/-}$ female crosses, the ratio of homozygous $Alkbh1^{-/-}$ pups was significantly skewed against females, with one female born for every three to four males (Fig. 3C, right panel). The survival of $Alkbh1^{-/-}$ male pups was 60% compared with $Alkbh1^{+/-}$ pups, whereas the proportion of viable $Alkbh1^{-/-}$ female pups was only 18%, $p=7.1\times10^{-5}$ (χ^2-test) (Fig. 3C, right panel). Following heterozygous crosses, the survival of $Alkbh1^{-/-}$ offspring was significantly reduced, 30% of the males and just 10% of the females survived compared with wild-type littermates, $p=1.4\times10^{-6}$ (χ^2-test (data not shown). A sex-ratio distortion was also seen in mid-stage $Alkbh1^{-/-}$ embryos (E10–E12.5) after heterozygous breedings (17 litters), with 89% male and 60% female embryos present compared with wild-type embryos (data not shown).

Spermatogenic Defects in *Alkbh1* Deficient Testis

Reduced testis weight was observed in $Alkbh1^{-/-}$ males at 12-week and 12-month of age, constituting three-quarters and two-thirds the mean weight of testis from wild-type littermates, respectively (Fig. 4A). TUNEL staining of testes from 12-week old wild-type and $Alkbh1^{-/-}$ males were histologically indistinguishable and showed no apoptotic cells (data not shown). On the other hand, extensive apoptosis and reduced number of germ cells were revealed in 5–10% of the seminiferous tubules in 9-month old $Alkbh1^{-/-}$ males (Fig. 4B, Fig. S2). In $Alkbh1^{-/-}$ testes, no apoptosis was detected in the spermatogonia (Sg) located at the edges of the tubules and in the meiotic spermatocytes (Sc) residing mostly in the two to three subbasal layers (Fig. 4B, Fig. S2). However, numerous apoptotic and degraded cells were seen in the subbasal regions corresponding to spermatocytes and spermatids, as well as in degenerating round and elongating spermatids (Sd) in the more luminal layers of the tubules (Fig. 4B, Fig. S2). In wild-type, a few apoptotic cells were occasionally located mainly at the basal layer of the seminiferous tubules (Fig. 4B, Fig. S2). To better define the basis for arrest in germ cells and the sex-ratio distortion, we focused on the XY-body in the pachytene stage of meiosis. The XY-body is a condensed chromatin structure containing the sex chromosomes, which is thought to be essential for meiotic progression. In mid-pachynema the XY-body forms a spherical structure near the nuclear periphery [25]. Two different markers against XY-bodies were used, macroH2A and FK2, however visible sex-bodies were readily identified in pachytene spermatocytes from 12-month old wild-type and $Alkbh1^{-/-}$ testes (Fig. 4C). MacroH2A recognizes the sex chromatin, and FK2 detects the abundant ubiquitination of H2A in the XY-body. We also did antibody staining against several specific stages throughout spermatogenesis, but no significant differences between wild-type and $Alkbh1^{-/-}$ mice were revealed (Fig. S3). The fact that sex-body formation is not impaired in *Alkbh1*-null males does not exclude the hypothesis of an epigenetic and silencing defect of the paternal X chromosome in those mice. Another possibility is that the skewing of the sexes in $Alkbh1^{-/-}$ mice is related to autosomal sex-ratio distortion. It is well known that most mechanisms that affect segregation distortion act in the male gametes and affect male fertility [24].

Expression Profiling in Wild-Type and $Alkbh1^{-/-}$ Testis

Due to the pivotal role of Alkbh1 in mouse survival and potentially in germ cells, we searched for Alkbh1-regulated genes

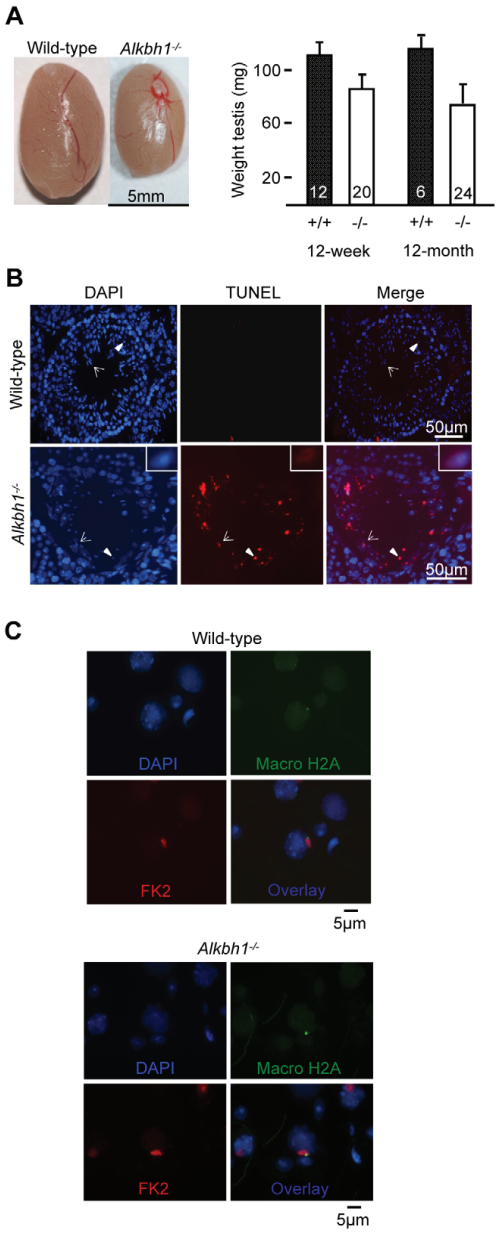


Figure 4. Spermatogenic defects in *Alkbh1* deficient testis. (A) Left panel, representative testis from 12-month old wild-type and *Alkbh1*$^{-/-}$ males. Right panel, average testis weight (mg) from 12-week old wild-type (110.6±10.4 mg, $n=12$) and *Alkbh1*$^{-/-}$ (85.6±8.4 mg, $n=20$) males, and 12-month old wild-type (114.4±8.9 mg, $n=6$) and *Alkbh1*$^{-/-}$ (74.3±12.1 mg, $n=24$) males. (B) TUNEL staining (middle panel) and DAPI staining (left panel) of testis sections from 9-month old wild-type and *Alkbh1*$^{-/-}$ mice, showing apoptosis in *Alkbh1*$^{-/-}$ round (arrowhead) and elongating (arrow) spermatids, and in degenerated germ cells in the subbasal layers of the tubules (middle and right panel). Closer view of one apoptotic elongating spermatid is shown in the lower panels. (C) Double immunostaining of XY-bodies in mid-pachytene cells. Testicular cells from 12-month old wild-type and *Alkbh1*$^{-/-}$ males were spread and stained with two different markers against XY-bodies. MacroH2A (green), FK2 (red) and DAPI (blue). (Magnifications: (B) ×20, (C) ×20).
doi:10.1371/journal.pone.0013827.g004

in adult testes. Microarray analysis of whole testes from 12-week old males identified 25 genes that were differentially expressed in *Alkbh1*$^{-/-}$ versus wild-type, using the class comparison strategy (Fig. 5A). *Ptpro* were also statistically significantly upregulated in *Alkbh1*$^{-/-}$ testes (Table S1; All data is deposited in GEO, accession number GSE22073). The function of PTPRO in adult testis has not been explored, but Avraham et al found expression of PTPRO in testis in humans [26]. Ptpro is suggested to be involved in the differentiation and axonogenesis of central and peripheral nervous system neurons, where it is in position to regulate phosphotyrosine levels in intracellular signaling cascades [27]. qPCR was performed on selected genes, to verify the class comparison analysis (Fig. 5B). Upregulation of *Vav2* and *Ccdc80* was confirmed in *Alkbh1*$^{-/-}$ versus wild-type whole testes. Vav2 is a guanine nucleotide exchange factor important for the formatin of adherens junctions between Sertoli cells and spermatids in testis, as well as in the formation of synapses in neurons [28]. The function of Ccdc80, also known as steroid sensitive gene 1, has not been studied in testis, but is supposed to be expressed in this organ according to its EST profile in the Unigene database (http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist = Mm.181074). Ccdc80 is expressed in human mesenchymal stem cells and mouse embryo cartilage, suggesting a role in skeletogenesis [29]. Together, these findings point towards a role in regulating the expression of genes having diverse functions – in spermatogenesis, in the nervous system and in skeletogenesis, although the genes affected in the microarray analysis are merely indirect targets of the Alkbh1 protein.

Alkbh1 Deficiency Causes Unilateral Eye Development

The reduced viability of *Alkbh1* deficient mice and the expression pattern of *Alkbh1* during embryonic development prompted us to analyse embryos and mice at earlier developmental stages. Both *Alkbh1*$^{+/-}$ and *Alkbh1*$^{-/-}$ mice showed embryonic (E) and postnatal (P) lethality, ranging from E9.5 to P28 (data not shown). Both embryos and neonatal mice clearly displayed an incompletely penetrant defect of small (microphthalmia) or missing (anophthalmia) eyes, and most often in the right eye (unilateral) (Fig. 6A, D). Eye malformations such as microphthalmia and anophthalmia occur in the mouse if eye morphogenesis is disrupted during the critical stages between E9.5 and E13.5 [30]. Small or missing eyes were observed in 18% of *Alkbh1*$^{-/-}$ embryos ($n=7/39$) and 9% of *Alkbh1*$^{+/-}$ embryos ($n=7/79$) at E11.5–E12.5. In surviving adults, eye defects were observed in 9% of *Alkbh1*$^{-/-}$ mice ($n=14/150$) and 0.5% of *Alkbh1*$^{+/-}$ mice ($n=1/198$). Eye defects varied from unilateral (one side) to bilateral (both sides) microphthalmia or anophthalmia, or unilateral microphthalmia in combination with unilateral anophthalmia (Fig. 6A, D). Intriguingly, the disturbed eye development affected the right eye more severely than the left eye, bearing resemblance to the histone arginine demethylase *Jmjd6* and the HMG box factor *Sox3* null phenotypes in mice [31–33].

To identify any abnormalities in addition to small or missing eyes, E18.5 embryos and newborn mice were analysed by skeletal

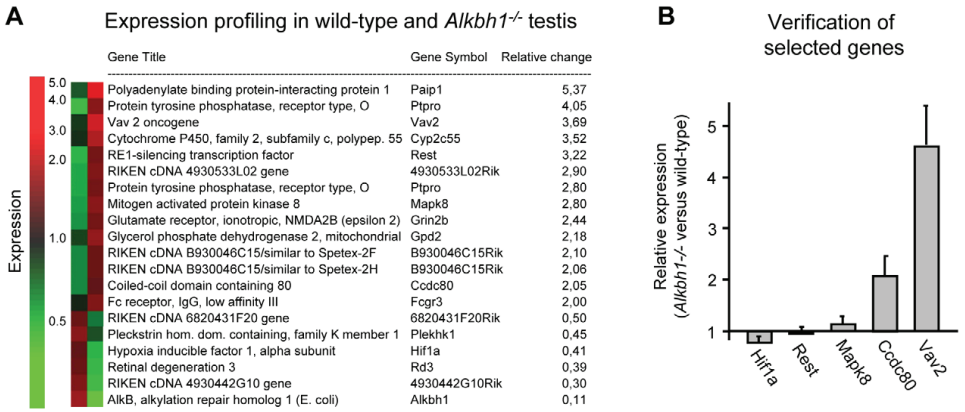


Figure 5. Expression profiling in wild-type and *Alkbh1*$^{-/-}$ adult testis. (A) Microarray analysis of whole testes from three wild-type and three *Alkbh1*$^{-/-}$ 12-week old males. Results are presented following class comparison analysis and visualized by GeneSpring v 6.0. (B) Verification of differentially expressed genes from the microarray analysis of wild-type and *Alkbh1*$^{-/-}$ testes. A selection of genes identified in the class comparison analysis (*Vav2, Mapk8, Ccdc80, Rest, Hif1a*) were checked for significance by qPCR. Upregulation of *Vav2* and *Ccdc80* were confirmed, while the differential expression of *Mapk8, Rest* and *Hif1a* were not found to be significant. On the RNA used for the microarray study. Reference sample, wild-type *Ccdc80* (RQ = 1.00); endogenous control, *18s*.
doi:10.1371/journal.pone.0013827.g005

staining of bone (Alizarin red) and cartilage (Alcian blue). Multiple defects were detected in the craniofacial, sternum and limb skeleton of mice lacking Alkbh1 (Figs. 6A–F). In the skull, reduced or missing intramembranous ossification resulted in enlarged sutures (Figs. 6B–C, F), while in the sternum, delayed ossification and aberrant fusion of the sternal bands were observed (Fig. 6E). Skeletal staining also showed assymetric shortening of the nasal bones, curving unilaterally in *Alkbh1*$^{-/-}$ mice causing mal-developed teeth (Fig. S4A), as well as reduced ossification in the phalanges (P) and the metatarsals (M) of the autopod of *Alkbh1*$^{-/-}$ newborns (Fig. S4B). The most crucial step in skeletal morphogenesis is the formation of mesenchymal condensations at E9.5 to E11.5 in mouse development [34]. The *Alkbh1* variable phenotype indicates incomplete condensation of mesenchymal cells during skeletogenesis.

Incomplete Penetrance of Unilateral Eye Defects

Penetrance is described as incomplete when a trait associated with a specific allele is expressed in a proportion of the population carrying the allele variant [35]. The eye phenotype associated with lack of the *Alkbh1* allele is characterized by incomplete penetrance (Fig. 7A–B). The *Alkbh1*$^{-/-}$ mouse in Fig. 7B has developed normally except for the deficiency of one eye. In contrast, the *Alkbh1*$^{-/-}$ embryo in Fig. 7A has gross developmental abnormalities, in addition to one small eye with only a residual mass of retinal cells, and one eye missing. The excessive brain tissue outside the skull is characteristic of a condition in which the neural tube fails to close, called exencephaly. Exencephaly is a neural tube defect (NTD), together with spina bifida (open spine) and anencephaly (open skull) [36]. At E10.5–E11.5, NTDs were observed in 23% of *Alkbh1*$^{+/-}$ embryos ($n = 12/52$) and 10% of *Alkbh1*$^{-/-}$ embryos ($n = 3/31$). The defects originated primarily from disrupted closure in the midbrain-hindbrain region (Fig. 7A) and upper spinal region, and were frequently associated with head and facial malformations (Fig. S4C). Around 50% of embryos with NTDs simultaneously displayed eye malformations ($n = 14/27$). The eye- and NTD-defects observed in *Alkbh1* mutants correspond with the expression pattern of *Alkbh1* seen in embryos and adult mice.

Gross morphological and histological analysis of adult *Alkbh1*$^{-/-}$ eyes revealed a range of serious deformities and size variations (Fig. 7C–D). Hematoxylin and eosin (HE) staining of paraffin-embedded sections showed that the lens was either completely missing or clearly smaller and displaced in the eye field (Fig. 7D). Furthermore, the lens fiber cells had lost their ordered lamination pattern, and swollen and liquefied fibers as well as vacuoles were seen throughout the lens (Fig. S4D). In retinal cells, there was a severe loss of organization even though all the retinal cells were present (Fig. 7D). In some areas, the neural retina (NR) was dysplastic with inclusions of rods and cones surrounded by outer nuclear layer cells (ONL), forming rosettes (Fig. 7E). In others, regions of thick layers of retinal pigment epithelium (RPE) cells were observed, with RPE cells appearing inside the NR layers in direct contact with the lens (Fig. 7E). Hence, Alkbh1 is important for growth and appropriate positioning and survival of lens and retinal cells.

Altered Expression of *Bmps* in *Alkbh1* Deficient Embryos

Embryonic development and tissue regeneration are regulated by four major families of signaling molecules. One of the largest families is the bone morphogenetic proteins (Bmps) [37]. In skeletogenesis, Bmp signaling plays an important role in regulating chondrocyte differentiation and establishment of joint boundaries [38]. Current evidence indicates that Bmp2, Bmp4 and Bmp7 are the main source of Bmp signaling in vertebrate limb buds [39]. Similar signaling mechanisms are suggested for growth and regional specification of the forebrain, branchial arches and eye during development [40–42]. This prompted us to examine the expression of *Bmp2*, *Bmp4* and *Bmp7* in apparently normal *Alkbh1*$^{-/-}$ embryos at E11.5 (Fig. 8A). *Bmp2* and *Bmp7* were induced in the lateral telencephalon (tc) of *Alkbh1*$^{-/-}$ embryos, and expression of *Bmp2* also increased in the frontonasal process (Fig. 8A). Moreover, *Bmp4* and *Bmp7* became upregulated specifically in the maxillary and mandibular cleft, while *Bmp2* was upregulated throughout the maxillary, mandibular and hyoid

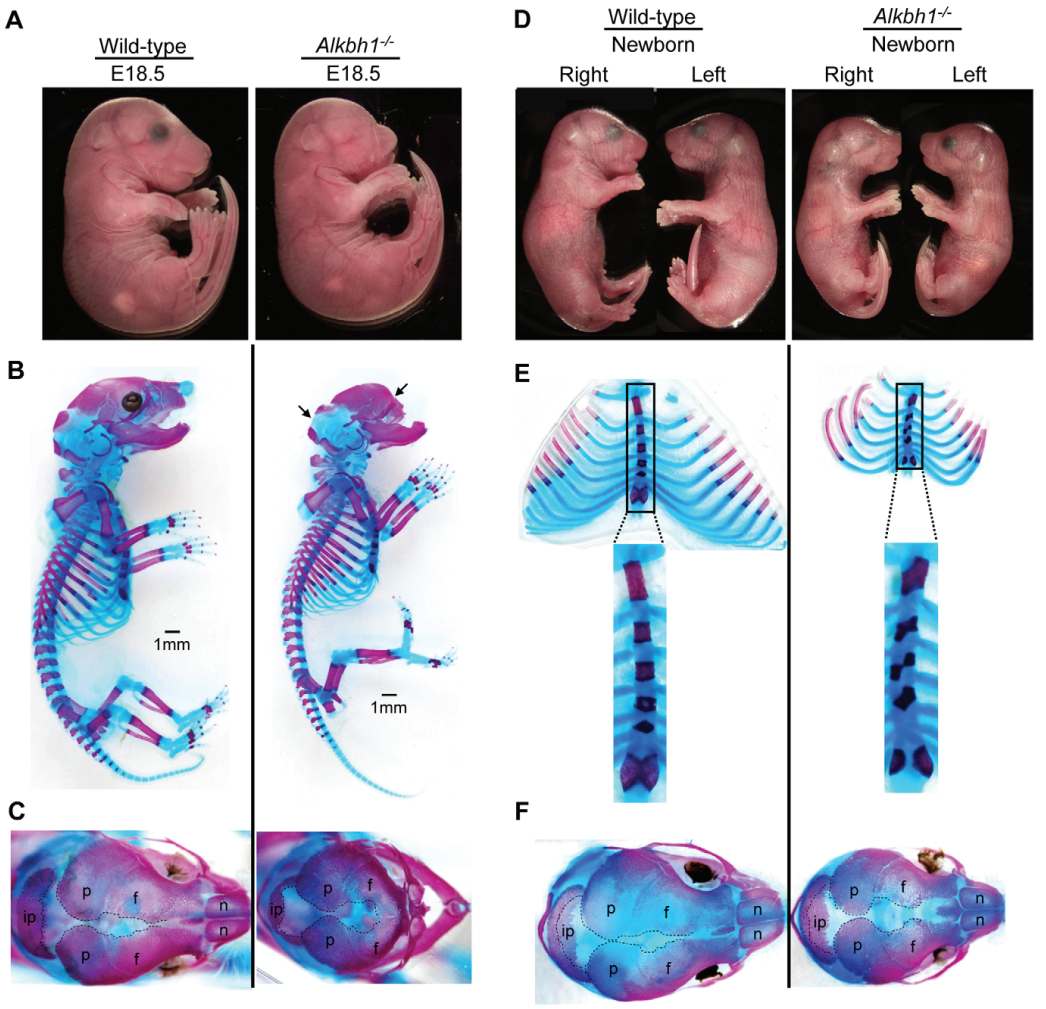


Figure 6. Eye and skeletal phenotype of *Alkbh1*$^{-/-}$ embryos and newborns, showing ossified areas in red and cartilage in blue. (A) Side view of embryos at E18.5. The *Alkbh1*$^{-/-}$ has a bilateral anophthalmic phenotype, a shortened snout and slightly reduced body size. (B) Skeletal staining of E18.5 embryos displaying missing nasal bones, shortened frontal bones and upward curving of the lower jaws in *Alkbh1*$^{-/-}$. (C) Dorsal view of the craniofacial skeleton of E18.5 embryos showing reduced ossification of the interparietal, parietal and frontal bones leading to bigger sutures in the skull of the *Alkbh1*$^{-/-}$. (D) Side view of newborn mice. The *Alkbh1*$^{-/-}$ has a unilateral microphthalmic eye phenotype and reduced body size. (E) Skeletal staining of the rib cage and sternum (in magnification) revealing delayed ossification and severe aberrant fusion of the sternal bands in *Alkbh1*$^{-/-}$ newborn mice. Interparietal, ip; parietal, p; frontal, f; nasal, n. (F) Dorsal view of the craniofacial skeleton of newborns demonstrating bigger sutures in the skull of *Alkbh1*$^{-/-}$.
doi:10.1371/journal.pone.0013827.g006

mesenchyme (Fig. 8A). In limb buds, *Bmp4* and *Bmp7* were highly upregulated in the apical ectodermal ridge (AER) and in two broader domains anteriorly and posteriorly (Fig. 8B). *Bmp2* expression disappeared from the posterior domain in hindlimb, and expression in AER of forelimb diffused proximally into the mesenchyme (lm) (Fig. 8B). The disrupted expression of *Bmp2*, *Bmp4* and *Bmp7* might be the cause of the somewhat smaller limb buds in *Alkbh1*$^{-/-}$ embryos. Regulation of these *Bmp* genes is important for AER formation, which is the major signaling center for limb outgrowth [37]. In general, both increased and decreased Bmp signaling can result in skeletal phenotypes [38].

Discussion

Our data point towards an important role of Alkbh1 in spermatogenesis and embryonic development. Several genes involved in spermatogenesis, in the nervous system and in skeletogenesis were found to be differentially expressed in

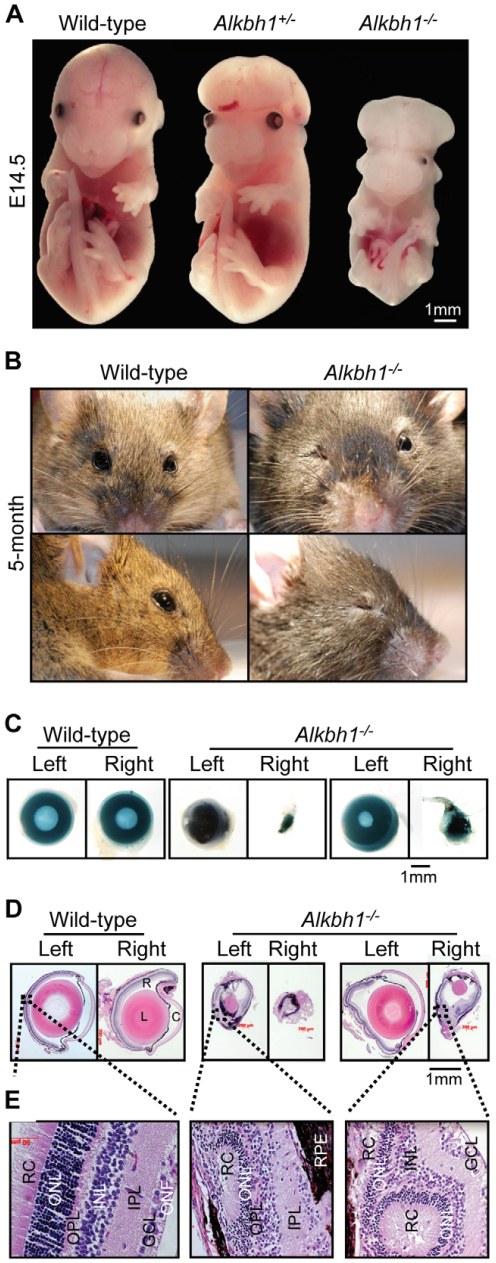


Figure 7. Incomplete penetrance of eye defects and exencephaly of *Alkbh1*$^{-/-}$ embryos and adults. (A) Frontal view of wild-type, *Alkbh1*$^{+/-}$, and *Alkbh1*$^{-/-}$ embryos at E14.5. The *Alkbh1*$^{+/-}$ and *Alkbh1*$^{-/-}$ embryos exhibit exencephaly in combination with a shortened, broad snout, while the *Alkbh1*$^{-/-}$ embryo also has a bilateral microphthalmic eye phenotype and severely reduced body size. (B) Frontal view and side view of wild-type and *Alkbh1*$^{-/-}$ adult mice. The *Alkbh1*$^{-/-}$ mouse has a unilateral microphthalmic eye malformation. (C) Whole-mount view of fixated eyes from wild-type and *Alkbh1*$^{-/-}$ adult mice demonstrating absent pupils and various degrees of eye malformations. (D) Histological analysis of paraffin-embedded eye sections. In *Alkbh1*$^{-/-}$ eyes, the lens is either missing or small and displaced in the eye field. Retinal cells appear degenerated or have lost their organized lamination pattern. R, retina; L, lens; C, cornea. (E) Closer view of the retina shown in d. Neural retinal cells are dysplastic with inclusions of rods and cones surrounded by outer nuclear layer cells. Retinal pigment epithelium cells are found inside the multi-layered neural retina. RC, rods and cones; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer; ONF, optic nerve fibers. (Magnifications: (D) ×2,5, (E) ×40).
doi:10.1371/journal.pone.0013827.g007

Alkbh1$^{-/-}$ whole testes. Adult males deficient in *Alkbh1* exhibited dramatically increased levels of apoptosis in 5–10% of the seminiferous tubules of testes; in spermatids and in degenerated germ cells in the subbasal regions corresponding to spermatocytes and spermatids. The reduced number of all spermatogenic cells in the apoptotic tubules, might reflect an indirect effect of prolonged arrest in spermatids in the affected tubules. Similar nonspecific defects have been seen in *miwi*-null mice [43] and TRF2 mutants [44,45]. Most genes involved in spermatogenesis display pleiotropic and leaky mutant phenotypes, as presented in this paper. Targeted disruptions of genes resulting in a variable range of defects and incomplete penetrance of spermatogenesis is even the case for regulatory genes, such as those encoding RNA binding proteins DAZLA [46] and MVH [47], and cell cycle regulators HSP-70.2 [48,49] and cyclin A1 [50].

The sex-ratio distortion lead us to study the XY-bodies in pachytene spermatocytes from *Alkbh1*$^{-/-}$ testes, however visible XY-bodies were detected showing that X and Y chromosomes paired normally during male meiosis. This does not exclude the hypothesis of an epigenetic and silencing defect of the paternal X chromosome in those mice, which could explain the sex-ratio distortion observed. Moreover, embryonic and postnathal lethality seen in *Alkbh1*$^{-/-}$ mice seem to be of paternal origin and *Alkbh1*$^{-/-}$ males exhibit subfertility compared to wild-type males. Several characteristics of the *Alkbh1*$^{-/-}$ mice are comparable with those described for the *Jmjd1a* histone lysine demethylase and the *G9a* histone lysine methyltransferase mutant mice, although to a milder extent than demonstrated in the histone disrupted mouse models [51,52]. *Jmjd1a* deficiency caused extensive germ cell apoptosis and blocked spermatid elongation, resulting in small testes and infertility in male mice [51]. Inactivation of *G9a* in the germ-lineage resulted in sterility due to a drastic loss of mature gametes [52]. The specific upregulation of *Alkbh1* in the pachytene stage, together with the sex-ratio distortion, suggests a potential to regulate the expression of genes during meiosis in the germline. Future investigations will focus on the regulation of specific genes in pachytene spermatocytes isolated from *Alkbh1*$^{-/-}$ and wild-type testes.

Alkbh1 mutant mice displayed phenotypes of incomplete penetrance, including unilateral eye malformations, neural tube defects, and craniofacial and skeleton associated abnormalities. Around 10% of the *Alkbh1*$^{-/-}$ mice appeared relatively normal, whereas the most affected mice died early during embryogenesis. The phenotypes are similar to published results on the bone morphogenetic proteins (Bmps), such as haploinsufficiency of Bmp2 causing exencephaly comparable to Fig. 7A [53], and compound heterozygous mutants for Bmp2 and Bmp4 showing unilateral microphthalmia similar to Fig. 6–7 [54]. In addition, postnatal lethality and sex-ratio distortion against females have been shown in *Bmp4*$^{tm1/+}$ heterozygous at weaning [55]. Altogether, this led us to investigate the effect on Bmps, and the

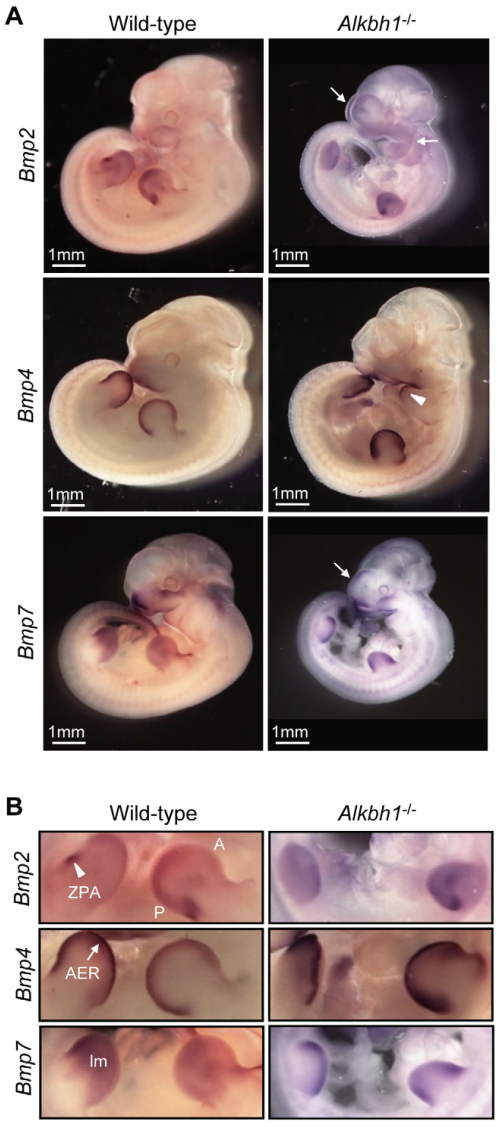


Figure 8. Misexpression of *Bmp2, Bmp4* and *Bmp7* in *Alkbh1* deleted embryos at E11.5. (A) Whole-mount *in situ* hybridization of *Bmp2, Bmp4* and *Bmp7* in *Alkbh1*$^{-/-}$ embryos at E11.5, side view. Altered expression is shown in the frontonasal process, the telencephalon (tc) and in the branchial arches (ba). (B) Closer view of the limbs from the whole-mount *in situ* hybridization of *Bmp2, Bmp4* and *Bmp7* in *Alkbh1*$^{-/-}$ embryos at E11.5, side view. The expression is altered in the zone of polarizing activity (ZPA), the apical ectodermal ridge (AER) and in the limb mesenchyme (lm). A, anterior; P, posterior.
doi:10.1371/journal.pone.0013827.g008

misexpression of *Bmp2*, *Bmp4* and *Bmp7* in *Alkbh1*$^{-/-}$ embryos at E11.5 might explain the inconsistent phenotypes presented. This is due to the critical dependence of gene dosage for proper Bmp function together with the expression- and function-overlap of the Bmps in different tissues [39,40]. Mouse models of Bmp4 and Bmp7 have shown that redundancy between Bmp4 and Bmp7 is not sufficient to prevent the eye phenotype to occur [40,41,56]. In the skull, signaling pathways involving Bmp2, Bmp4 and Bmp7 regulate mesenchymal condensation size, and intense expression of these signaling genes is necessary for closure of sutures [34]. In addition to modifier genes such as Bmps, genetic and epigenetic components can cause variable phenotypic outcomes from specific genes [57], leading to irregular patterns of inheritance as seen for the *Alkbh1* deficient mice. A recent paper has shown that the osteoblast-specific transcription factor Osterix is regulated by the JmjC histone demethylase NO66 [58]. Experiments in the chick embryo have revealed that epigenetic factors are required for the establishment of left-right asymmetries, together with the action of well-studied genetic and signaling mechanisms [59,60].

The reduced viability and developmental phenotypes apparent in our mouse model, was not reported in the *Alkbh1*-null mice generated by Pan et al [20]. However, they showed severe growth defects in *Alkbh1*$^{-/-}$ embryos and newborns in addition to placentas [20], and the growth retardation demonstrated in pups at four weeks of age are comparable with our data (Pan et al. Suppl. Fig. 2 and this paper Fig. S1). No obvious color variation (from red/pink to pale brown/bluish) or growth retardation was observed in *Alkbh1*$^{-/-}$ placentas compared to wild-type placentas. Our results are based on extensive breeding studies of *Alkbh1* targeted mice, which revealed a dramatic effect on lethality and sex-ratio in adult mice. We therefore sought to characterize testes and embryos in more detail, as well as the prominent abnormalities in eye development. The different mouse background chosen as well as the dissimilar targeting strategies deleting different parts of the *Alkbh1* gene (Exon 6 in our strain, Exon 3 in Pan et al) could be a possible explanation for the discrepancies in the penetrance of phenotypes in the two knockout mice models. Even so, together with the findings on Alkbh1 by Pan et al, these data suggest that the effect of *Alkbh1* deficiency is pleiotropic and dependent on cell type and/or stage of development.

Recent studies have recognized roles for 2-oxoglutarate dependent dioxygenases in histone and nucleic acid demethylation, as well as in signaling protein hydroxylation [19]. For the demethylating enzymes, several have been shown to carry out its reaction in a manner similar to the potential Alkbh1 mediated, iron- and 2-oxoglutarate dependent, hydroxylation [1,2,61]. Previously, mouse models for histone methyl transferases and histone demethylases have been characterized with multiple developmental defects [31–33,62]. Our working hypothesis, based on the variable developmental phenotype of *Alkbh1* deficient mice together with the localization of Alkbh1 to nuclear euchromatin [20], is that Alkbh1 possibly works as a histone demethylase during embryogenesis and spermatogenesis. We believe that the hydroxylation activity of Alkbh1 is dependent on yet undefined partners specific for the different stages/tissues where it has an important role, and this will be addressed in future studies for the pachytene stage of meiosis in male germ cells – when homologues chromosomes pair and crossing over can occur.

Supporting Information

Figure S1 Average body weight of *Alkbh1* targeted males and females. (A) 1-month old wild-type (19.0±2.0 g, $n = 45$) and *Alkbh1*$^{-/-}$ (14.6±3.8 g, $n = 48$) males, and 1-month old wild-type (17.7±1.7 g, $n = 43$) and *Alkbh1*$^{-/-}$ (14.8±2.2 g, $n = 33$) females. The average weight was 25% lower for *Alkbh1*$^{-/-}$ males than for wild-type males and 15% lower for *Alkbh1*$^{-/-}$ females than for wild-

type females. About one out of five *Alkbh1*$^{-/-}$ males showed more than 40% lower weight compared to wild-type males. (B) 9-month old wild-type (40.5±4.2 g, $n = 23$) and *Alkbh1*$^{-/-}$ (32.5±2.9 g, $n = 31$) males, and 9-month old wild-type (31.5±3.4 g, $n = 28$) and *Alkbh1*$^{-/-}$ (29.3±3.8 g, $n = 41$) females. The average weight of *Alkbh1*$^{-/-}$ males was 20% below that of wild-type males, and the average weight of *Alkbh1*$^{-/-}$ females was 7% below that of wild-type females. No weight difference was demonstrated between the *Alkbh1*$^{+/-}$ and wild-type (data not shown). +/+ (wild-type), black bars; −/− (*Alkbh1*$^{-/-}$), grey bars.
Found at: doi:10.1371/journal.pone.0013827.s001 (0.10 MB TIF)

Figure S2 Closer view of the DAPI and TUNEL staining of testis sections shown in Fig. 4. (A, B) Sections from 9-month old wild-type (left panel) and *Alkbh1*$^{-/-}$ (right panel) mice are presented. Apoptosis was detected in degenerating spermatids (Sd) in the luminal layers of *Alkbh1*$^{-/-}$ tubules, as well as in severely degraded cells in the subbasal regions corresponding to spermatocytes and spermatids. No apoptotic cells were seen in spermatogonia (Sg) and spermatocytes (Sc) in *Alkbh1*$^{-/-}$ mice, although the amount of all spermatogenic cells are reduced in the apoptotic tubules. (Magnification: ×20).
Found at: doi:10.1371/journal.pone.0013827.s002 (3.93 MB TIF)

Figure S3 Immunostaining with stage-specific antibodies against spermatogenic cells in *Alkbh1* deficient testes. (A) Testis sections from 12-month old wild-type and *Alkbh1*$^{-/-}$ males stained with TRA98 antibody specific for spermatogonia, which were present both in wild-type and mutant. Although several tubules showed spermatogonia not only in the first basal layer, but also in the subbasal layers in the *Alkbh1*$^{-/-}$ mice, no significant differences were detected when compared to wild-type. (B) Testis sections from 12-month old wild-type and *Alkbh1*$^{-/-}$ males stained with TRA369 specific for pachytene spermatocytes through elongating spermatids, which were present both in wild-type and mutant. (Magnification: ×20).
Found at: doi:10.1371/journal.pone.0013827.s003 (1.84 MB TIF)

Figure S4 Skeletal defects, eye defects in combination with NTD, and lens defects in *Alkbh1* targeted mice. (A) Craniofacial defects. Dorsal view of the craniofacial skeleton of adult mice showing assymetric shortening of the nasal bones, curving unilaterally in *Alkbh1*$^{-/-}$ mice causing mal-developed teeth ($n = 4$ *Alkbh1*$^{-/-}$; $n = 1$ *Alkbh1*$^{+/-}$). Ossified areas are shown in red and cartilage in blue. (B) Limb defects. Dorsal view of the autopod limb skeleton revealing reduced ossification in the phalanges (P) and the metatarsals (M) of the autopod of *Alkbh1*$^{-/-}$ newborns ($n = 4/4$ *Alkbh1*$^{-/-}$). Ossified areas are shown in black and cartilage in blue. (C) Eye defects and NTDs. Side view of embryos at E12.5. The *Alkbh1*$^{-/-}$ embryo has a bilateral microphthalmic eye phenotype in combination with a neural tube defect (NTD). The NTD is originating from disrupted closure in the upper spinal region, and is associated with head and facial malformations leading to a shortened, broad snout. In addition, a severe intracranial hemorrhage is visible. (D) Lens defects. Histological analysis of paraffin-embedded eye sections from adult mice. In *Alkbh1*$^{-/-}$ eyes the lens fiber cells have lost their ordered lamination pattern, and swollen and liquefied fibers as well as vacuoles are seen throughout the lens. (Magnification: ×10).
Found at: doi:10.1371/journal.pone.0013827.s004 (6.28 MB TIF)

Table S1 Statistically upregulated genes in *Alkbh1*$^{-/-}$ versus wild-type testes identified in the microarray analysis. Microarray analysis of RNA extracted from whole testes from three wild-type and three *Alkbh1*$^{-/-}$ 12-week old males identified 6 genes that were statistically upregulated in *Alkbh1*$^{-/-}$ versus wild-type. To find differentially expressed genes, t-test with randomized variance was used as statistical test and the cut-off (p-value) was set to 0.05 with a FDR correction.
Found at: doi:10.1371/journal.pone.0013827.s005 (0.22 MB TIF)

Acknowledgments

We are grateful to Hege Wiksèn, Cecilie G. Castellanos, Linda Ellevog and Gaute Nesse for excellent technical assistance. We thank IngenKO, Australia, The Norwegian Transgenic Center (NTS) and the Centre for Comparative Medicine at Oslo University Hospital for the excellent service they provided.

Author Contributions

Conceived and designed the experiments: LMN JTL AK. Performed the experiments: LMN JS EL AN RO KF GFL. Analyzed the data: LMN JS EL AN RO KF TR SHN JTL AK. Wrote the paper: LMN AK.

References

1. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 419: 174–178.
2. Falnes PO, Johansen RF, Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature 419: 178–182.
3. Kurowski MA, Bhagwat AS, Papaj G, Bujnicki JM (2003) Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. BMC Genomics 4: 48.
4. Falnes PO, Klungland A, Alseth I (2007) Repair of methyl lesions in DNA and RNA by oxidative demethylation. Neuroscience 145: 1222–1232.
5. Lee DH, Jin SG, Cai S, Chen Y, Pfeifer GP, et al. (2005) Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. J Biol Chem 280: 39448–39459.
6. Duncan T, Trewick SC, Koivisto P, Bates PA, Lindahl T, et al. (2002) Reversal of DNA alkylation damage by two human dioxygenases. Proc Natl Acad Sci U S A 99: 16660–16665.
7. Aas PA, Otterlei M, Falnes PO, Vagbo CB, Skorpen F, et al. (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421: 859–863.
8. Sedgwick B (2004) Repairing DNA-methylation damage. Nat Rev Mol Cell Biol 5: 148–157.
9. Monsen VT, Sundheim O, Aas PA, Westbye MP, Sousa MM, et al. (2010) Divergent {beta}-hairpins determine double-strand versus single-strand substrate recognition of human AlkB-homologues 2 and 3. Nucleic Acids Res, gkq518 [pii];10.1093/nar/gkq518 [doi].
10. Ringvoll J, Nordstrand LM, Vagbo CB, Talstad V, Reite K, et al. (2006) Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. EMBO J 25: 2189–2198.
11. Fu D, Brophy JA, Chan CT, Atmore KA, Begley U, et al. (2010) Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. Mol Cell Biol 30: 2449–2459.
12. Songe-Moller L, van den BE, Leihne V, Vagbo CB, Kristoffersen T, et al. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. Mol Cell Biol 30: 1814–1827.
13. Shimada K, Nakamura M, Anai S, De VM, Tanaka M, et al. (2009) A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. Cancer Res 69: 3157–3164. 0008-5472.CAN-08-3530 [pii];10.1158/0008-5472.CAN-08-3530 [doi].
14. Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, et al. (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 318: 1469–1472.
15. Boissel S, Reish O, Proulx K, Kawagoe-Takaki H, Sedgwick B, et al. (2009) Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet 85: 106–111. S0002-9297(09)00238-9 [pii];10.1016/j.ajhg.2009.06.002 [doi].
16. Han Z, Niu T, Chang J, Lei X, Zhao M, et al. (2010) Crystal structure of the FTO protein reveals basis for its substrate specificity. Nature 464: 1205–1209. nature08921 [pii];10.1038/nature08921 [doi].
17. Westbye MP, Feyzi E, Aas PA, Vagbo CB, Talstad VA, et al. (2008) Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J Biol Chem 283: 25046–25056.
18. Sedgwick B, Robins P, Lindahl T (2006) Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases. Methods Enzymol 408: 108–120.
19. Loenarz C, Schofield CJ (2008) Expanding chemical biology of 2-oxoglutarate oxygenases. Nat Chem Biol 4: 152–156.

20. Pan Z, Sikandar S, Witherspoon M, Dizon D, Nguyen T, et al. (2008) Impaired placental trophoblast lineage differentiation in Alkbh1(−/−) mice. Dev Dyn 237: 316–327.
21. Tsujikawa K, Koike K, Kitae K, Shinkawa A, Arima H, et al. (2007) Expression and sub-cellular localization of human ABH family molecules. J Cell Mol Med 11: 1105–1116.
22. Bellve AR (1993) Purification, culture, and fractionation of spermatogenic cells. Methods Enzymol 225: 84–113.
23. Kataoka H, Yamamoto Y, Sekiguchi M (1983) A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate. J Bacteriol 153: 1301–1307.
24. Taylor DR, Ingvarsson PK (2003) Common features of segregation distortion in plants and animals. Genetica 117: 27–35.
25. Solari AJ (1974) The behavior of the XY pair in mammals. Int Rev Cytol 38: 273–317.
26. Avraham S, London R, Tulloch GA, Ellis M, Fu Y, et al. (1997) Characterization and chromosomal localization of PTPRO, a novel receptor protein tyrosine phosphatase, expressed in hematopoietic stem cells. Gene 204: 5–16.
27. Beltran PJ, Bixby JL, Masters BA (2003) Expression of PTPRO during mouse development suggests involvement in axonogenesis and differentiation of NT-3 and NGF-dependent neurons. J Comp Neurol 456: 384–395.
28. Kawakatsu T, Ogita H, Fukuhara T, Fukuyama T, Minami Y, et al. (2005) Vav2 as a Rac-GDP/GTP exchange factor responsible for the nectin-induced, c-Src- and Cdc42-mediated activation of Rac. J Biol Chem 280: 4940–4947.
29. Liu Y, Monticone M, Tonachini L, Mastrogiacomo M, Marigo V, et al. (2004) URB expression in human bone marrow stromal cells and during mouse development. Biochem Biophys Res Commun 322: 497–507.
30. Graw J (2003) The genetic and molecular basis of congenital eye defects. Nat Rev Genet 4: 876–888.
31. Rizzoti K, Lovell-Badge R (2007) SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. Development 134: 3437–3448.
32. Chang B, Chen Y, Zhao Y, Bruick RK (2007) JMJD6 is a histone arginine demethylase. Science 318: 444–447.
33. Bose J, Gruber AD, Helming L, Schiebe S, Wegener I, et al. (2004) The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. J Biol 3: 15.
34. Hall BK, Miyake T (2000) All for one and one for all: condensations and the initiation of skeletal development. Bioessays 22: 138–147.
35. Glazier AM, Nadeau JH, Aitman TJ (2002) Finding genes that underlie complex traits. Science 298: 2345–2349.
36. Copp AJ, Greene ND, Murdoch JN (2003) The genetic basis of mammalian neurulation. Nat Rev Genet 4: 784–793.
37. Robert B (2007) Bone morphogenetic protein signaling in limb outgrowth and patterning. Dev Growth Differ 49: 455–468.
38. Baldridge D, Shchelochkov O, Kelley B, Lee B (2010) Signaling pathways in human skeletal dysplasias. Annu Rev Genomics Hum Genet 11: 189–217. 10.1146/annurev-genom-082908-150158 [doi].
39. Bandyopadhyay A, Tsuji K, Cox K, Harfe BD, Rosen V, et al. (2006) Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis. PLoS Genet 2: e216.
40. Ducy P, Karsenty G (2000) The family of bone morphogenetic proteins. Kidney Int 57: 2207–2214.
41. Ohkubo Y, Chiang C, Rubenstein JL (2002) Coordinate regulation and synergistic actions of BMP4, SHH and FGF8 in the rostral prosencephalon regulate morphogenesis of the telencephalic and optic vesicles. Neuroscience 111: 1–17.
42. Wordinger RJ, Clark AF (2007) Bone morphogenetic proteins and their receptors in the eye. Exp Biol Med (Maywood) 232: 979–992. 232/8/979 [pii];10.3181/0510-MR-345 [doi].
43. Deng W, Lin H (2002) miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell 2: 819–830.
44. Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P, et al. (2001) Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. Mol Cell 7: 509–515.
45. Zhang D, Penttila TL, Morris PL, Teichmann M, Roeder RG (2001) Spermiogenesis deficiency in mice lacking the Trf2 gene. Science 292: 1153–1155.
46. Ruggiu M, Speed R, Taggart M, McKay SJ, Kilanowski F, et al. (1997) The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. Nature 389: 73–77.
47. Tanaka SS, Toyooka Y, Akasu R, Katoh-Fukui Y, Nakahara Y, et al. (2000) The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes Dev 14: 841–853.
48. Dix DJ, Allen JW, Collins BW, Mori C, Nakamura N, et al. (1996) Targeted gene disruption of Hsp70-2 results in failed meiosis, germ cell apoptosis, and male infertility. Proc Natl Acad Sci U S A 93: 3264–3268.
49. Zhu D, Dix DJ, Eddy EM (1997) HSP70-2 is required for CDC2 kinase activity in meiosis I of mouse spermatocytes. Development 124: 3007–3014.
50. Liu D, Matzuk MM, Sung WK, Guo Q, Wang P, et al. (1998) Cyclin A1 is required for meiosis in the male mouse. Nat Genet 20: 377–380.
51. Liu Z, Zhou S, Liao L, Chen X, Meistrich M, et al. (2009) The Jmjd1a demethylase-regulated histone modification is essential for crem-regulated gene expression and spermatogenesis. J Biol Chem.
52. Tachibana M, Nozaki M, Takeda N, Shinkai Y (2007) Functional dynamics of H3K9 methylation during meiotic prophase progression. EMBO J 26: 3346–3359.
53. Castranio T, Mishina Y (2009) Bmp2 is required for cephalic neural tube closure in the mouse. Dev Dyn 238: 110–122.
54. Uchimura T, Komatsu Y, Tanaka M, McCann KL, Mishina Y (2009) Bmp2 and Bmp4 genetically interact to support multiple aspects of mouse development including functional heart development. Genesis 47: 374–384.
55. Dunn NR, Winnier GE, Hargett LK, Schrick JJ, Fogo AB, et al. (1997) Haploinsufficient phenotypes in Bmp4 heterozygous null mice and modification by mutations in Gli3 and Alx4. Dev Biol 188: 235–247.
56. Wyatt AW, Osborne RJ, Stewart H, Ragge NK (2010) Bone morphogenetic protein 7 (BMP7) mutations are associated with variable ocular, brain, ear, palate, and skeletal anomalies. Hum Mutat 31: 781–787. 10.1002/humu.21280 [doi].
57. van Heyningen V, Yeyati PL (2004) Mechanisms of non-Mendelian inheritance in genetic disease. Hum Mol Genet 13 Spec No 2: R225–R233.
58. Sinha KM, Yasuda H, Coombes MM, Dent SY, de CB (2010) Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. EMBO J 29: 68–79. emboj2009332 [pii];10.1038/emboj.2009.332 [doi].
59. Raya A, Kawakami Y, Rodriguez-Esteban C, Ibanes M, Rasskin-Gutman D, et al. (2004) Notch activity acts as a sensor for extracellular calcium during vertebrate left-right determination. Nature 427: 121–128.
60. Wang S, Yu X, Zhang T, Zhang X, Zhang Z, et al. (2004) Chick Pcl2 regulates the left-right asymmetry by repressing Shh expression in Hensen's node. Development 131: 4381–4391.
61. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, et al. (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439: 811–816.
62. Li MO, Sarkisian MR, Mehal WZ, Rakic P, Flavell RA (2003) Phosphatidylserine receptor is required for clearance of apoptotic cells. Science 302: 1560–1563.

II

Nucleic Acids Research Advance Access published September 10, 2012

Nucleic Acids Research, 2012, 1–14
doi:10.1093/nar/gks839

Alkbh1 and Tzfp repress a non-repeat piRNA cluster in pachytene spermatocytes

Line M. Nordstrand[1], Kari Furu[1], Jonas Paulsen[2], Torbjørn Rognes[1,3] and Arne Klungland[1,4,*]

[1]Department of Microbiology, Centre for Molecular Biology and Neuroscience, Oslo University Hospital, Rikshospitalet, PO Box 4950, [2]Institute for Medical Informatics, Oslo University Hospital, The Norwegian Radium Hospital, PO Box 4953, Nydalen, NO-0424 Oslo, [3]Department of Informatics, University of Oslo, PO Box 1080, Blindern, NO-0316 Oslo and [4]Institute of Basic Medical Sciences, University of Oslo, PO Box 1018, Blindern, NO-0315 Oslo, Norway

Received July 4, 2012; Revised August 9, 2012; Accepted August 10, 2012

ABSTRACT

Piwi proteins and Piwi-interacting small RNAs (piRNAs) have known functions in transposon silencing in the male germline of fetal and newborn mice. Both are also present in adult testes; however, their function here remains a mystery. Here, we confirm that most piRNAs in meiotic spermatocytes originate from clusters in non-repeat intergenic regions of DNA. The regulation of these piRNA clusters, including the processing of the precursor transcripts into individual piRNAs, is accomplished through mostly unknown processes. We present a possible regulatory mechanism for one such cluster, named cluster 1082B, located on chromosome 7 in the mouse genome. The 1082B precursor transcript and its 788 unique piRNAs are repressed by the Alkbh1 dioxygenase and the testis-specific transcription repressor Tzfp. We observe a remarkable >1000-fold upregulation of individual piRNAs in pachytene spermatocytes isolated from *Alkbh1*- and *Tzfp*-deficient murine testes. Repression of cluster 1082B is further supported by the identification of a 10-bp Tzfp recognition sequence contained within the precursor transcript. Downregulation of LINE1 and IAP transcripts in the *Alkbh1*- and *Tzfp*-deficient mice leads us to propose a potential role for the 1082B-encoded piRNAs in transposon control.

INTRODUCTION

Spermatogenesis is a cyclic developmental process by which spermatogonia generate mature spermatozoa (1,2). The process takes place in the seminiferous tubules of the testis and involves three phases: mitotic proliferation of spermatogonia, meiosis and finally, spermiogenesis, which involves a stepwise maturation of the spermatids to mature spermatozoa. Meiosis involves two successive cell divisions of spermatocytes resulting in the formation of haploid spermatids. During the prolonged prophase I of meiosis, the primary spermatocytes go through the leptotene, zygotene, pachytene, diplotene and diakinesis stages in a highly organized, sequential manner that involves homologous chromosome pairing, synaptonemal complex formation and meiotic recombination.

Piwi-interacting small RNA (piRNA) is a distinct class of small non-coding RNAs. They are defined by their association with Piwi proteins, a subgroup of the Argonaute protein family (AGO), which are well-known mediators in small RNA-mediated silencing (3). The Piwi proteins were first identified in *Drosophila melanogaster* (4) but have now been characterized in many different organisms, including mammals. Three murine Piwi proteins have been identified: Miwi, Mili and Miwi2 (5–7). Unlike other small non-coding RNAs like microRNAs (miRNAs) and endo-small interfering RNAs (endo-siRNAs), piRNAs seem to be highly specific for the male germline in mammals (8–11). All three Piwi proteins are essential for spermatogenesis in mouse as deletion of all the genes renders male knockout mice infertile (5–7).

*To whom correspondence should be addressed. Tel: +47 23074072; Fax: +47 23074061; Email: arne.klungland@rr-research.no
Present address:
Arne Klungland, Department of Microbiology, Centre for Molecular Biology and Neuroscience, Oslo University Hospital, Rikshospitalet, PO Box 4950, Nydalen, NO-0424 Oslo, Norway.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Each of the three mouse Piwi proteins associate with a specific subset of piRNAs and have different expression patterns (12). Mili and Miwi2 are expressed during embryogenesis and just after birth (7,8) and bind piRNAs called pre-pachytene piRNAs. Deep sequencing analyses have shown that these piRNAs often map to repeat-associated DNA sequences (13), and it is well established that these complexes are important for transposon silencing (7,14–16). Mili is also expressed in adults and can be detected in all spermatogenic cells until the round spermatid stage (8). Miwi is expressed in adult animals and can be detected in pachytene spermatocytes to elongating spermatids (5). Miwi-associated piRNAs are called pachytene piRNAs and are relatively depleted of repeat elements (13). Whether the Piwi–piRNA complexes in adult testes primarily control RNA stability, gene transcription, chromatin organization or protein synthesis is not well understood (17). However, both Mili and Miwi have been shown to be endonucleases (15,18,19), and their function is thought, at least in part, to be piRNA-guided degradation of target transcripts. In addition, some evidence suggest that they may be involved in translational control (20,21).

piRNAs are believed to derive from long single-stranded transcripts and processed either through a primary processing pathway or the so-called ping-pong amplification cycle (15,22,23). The production of pachytene piRNAs from long primary transcripts was recently confirmed by Ragan *et al*. using the novel NORAHDESK tool (24). The primary processing probably involves splicing of long transcripts from piRNA-rich genomic sequences called piRNA clusters (8,9). This process is thought to be independent of the endonuclease activity of Piwi proteins (25). piRNA clusters may extend for tens of thousands of bases, and each cluster encodes a precursor transcript that can generate many different piRNA sequences, some of which may be partially overlapping. No secondary structures for the primary transcripts have been determined. Many clusters give rise to piRNAs that map to both genomic strands, suggesting bidirectional transcription (13,15,22).

We previously characterized the *Alkbh1*$^{-/-}$ mouse, which display developmental defects and sex-ratio distortion. Mutant animals are viable and fertile, but have a decreased survival rate and display increased apoptosis in adult testes (26). Alkbh1 is a member of the mammalian AlkB family of dioxygenases and is proposed to be involved in epigenetic regulation (26–29). Alkbh1 localizes to nuclear euchromatin (27), and recent studies suggest that ALKBH1 is a histone H2A dioxygenase (Ougland *et al*., submitted for publication). Recently, we generated a *Tzfp* mouse model (hereafter referred to as the *Tzfp*$^{GTi/GTi}$ mouse). Tzfp is a testis-specific transcription repressor belonging to the BTB/POZ Zn finger family. Mice lacking this protein are viable and fertile and have no obvious phenotype (unpublished data). The zinc finger domain of Tzfp binds to a specific genomic sequence, the Tzfp-binding site (tbs) (TGTACAGTGT) located upstream of the *Aie1* gene (30). The interaction with the tbs has a repressive effect on the target gene. Expression analyses have shown that both *Tzfp* and *Alkbh1* are highly expressed in adult testes and that the expression peaks at the pachytene stage during spermatogenesis (26) (Furu *et al*., in preparation).

In the present work, we present a possible regulatory mechanism of a non-repeat piRNA cluster located on chromosome 7, consisting of 788 unique anti-sense piRNAs. The precursor transcript and in-dividual piRNAs derived from this cluster are dramatically upregulated in the absence of Alkbh1 and Tzfp, as shown in the two mouse models *Tzfp*$^{GTi/GTi}$ and *Alkbh1*$^{-/-}$. The cluster contains the Tzfp-binding sequence, implying that the interaction partners Tzfp and Alkbh1 regulate the expression of the precursor by binding to the tbs motif. Further, we provide evidence for repression of long interspersed elements 1 (LINE1) and intracisternal A-particle (IAP) transcripts in these mouse models, indicating that upregulated piRNAs derived from cluster 1082B are involved in transposon control.

MATERIALS AND METHODS

Mouse handling and genotyping

All mice experiments were approved by the Norwegian Animal Research Authority (Ref. nr. 08/9940; 12/10-2388) and done in accordance with institutional guidelines at the Centre for Comparative Medicine at Oslo University Hospital. Animal work was conducted in accordance with the rules and regulations of the Federation of European Laboratory Animal Science Associations. *Alkbh1*-null mice were previously generated by targeted deletion of exon 6, and heterozygous were backcrossed onto a C57BL6/J background (26). *Tzfp*-targeted mice were designed by a gene-trap procedure using an embryonic stem (ES) cell clone containing a splice acceptor site from the gene trap vector (Omnibank Gene Trap Vector 76) upstream of exon 1 of the *Tzfp* gene (Furu and Klungland, in preparation). The clone was obtained from Texas Institute of Genomic Medicine C57BL6/J ES cell clone library. For genotyping, ear-clip samples were degraded by incubation in 75 µl Hot Shot Lysis Buffer (25 mM NaOH, 0.2 mM Na2EDTA, pH 12) at 95°C for 30 min and then cooled down to 4°C before adding 75 µl Hot Shot Neutralization Buffer (40 mM Tris–HCl, pH 5). Samples were PCR amplified for 35 cycles with an annealing temperature of 60°C for the *Alkbh1* gene and 40 cycles with an annealing temperature of 58°C for the *Tzfp* gene. See primers 1–6 in Supplementary Table S1.

DNA microarray analysis

Total RNA was isolated from three *Alkbh1*$^{+/+}$ and three *Alkbh1*$^{-/-}$ 12-week-old testes using the Fast RNA Pro Green Kit (MP Biomedicals) according to the manufacturer's protocol. DNA remnants were removed using TURBO DNase (Ambion), and the RNA quality was checked using Agilent Bioanalyzer 2100 (RIN value between 9.8 and 10.0). Fifteen micrograms of biotinylated and fragmented cRNA was then hybridized onto the GeneChip Mouse Genome 430 2.0 Array (Affymetrix) according to the manufacturer's protocol.

Quality checks including scale factor, background, noise, spikes and RNA degradation were performed and validated using the yaqcaffy library (http://www.bioconductor.org/packages/2.3/bioc/%20vignettes/yaqcaffy/inst/doc/yaqcaffy.pdf). Affymetrix raw data were generated with GCOS 1.4 (GeneChip Operating Software, Affymetrix), and the signal intensities of each probe set were normalized with the robust microarray analysis (RMA) algorithm. To find differentially expressed genes, *t*-test with randomized variance was used as statistical test, and the cutoff (*P* value) was set to 0.05 with a FDR correction. Fold change for all genes that passed the above criteria was computed, and only the genes with ≥2-fold change were studied. A heatmap was generated using the GeneSpring GX 10 demoversion (Agilent). All data are MIAME compliant, and the raw data have been deposited in a MIAME compliant database under accession number GSE22073.

STAPUT isolation of pachytene cells

Pachytene cells were isolated from C57BL6/J, *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ 12-week-old testes using an adapted version of the STAPUT method (Bellvè, 1993). Each isolation required a total of six males, and 2× six mice for each genotype were used. The mice were killed using CO_2 and testes were taken out and put in ice-cold DMEM medium containing antibiotics. Testes were detunicated and the tubules treated with DNaseI, collagenase, trypsin and hyaluronidase (all from Sigma-Aldrich) at 34°C to remove connective tissue and somatic cells, yielding a cell suspension of germinal cells in DMEM containing 0.5% BSA. The cell suspension was loaded into the cell loading chamber of the STAPUT apparatus and separated by sedimentation velocity at unit gravity in a 2–4% w/v BSA gradient in DMEM medium at 4°C for 2.5 h. After sedimentation, 10-ml fractions were collected and checked under the microscope. Fractions containing pachytene spermatocytes were pooled and the cell number was counted in a Countess® Automated Cell Counter (Invitrogen). Cells were spun down, and the pellet was snap frozen in liquid nitrogen before being placed in −70°C. An aliquot of purified cells was fixed on SuperFrost Plus slides (VWR) using cell adherence solution (Crystalgen, Lot no 425081) for microscopic analyses (Supplementary Figure S2). One isolation yielded $\sim 1.5 \times 10^6$ pachytene cells with an average size of 12.5 µm.

TaqMan® and SYBR® green gene expression analysis

Total RNA was isolated from C57BL6/J, *Tzfp*$^{GTi/GTi}$ and *Alkbh1*$^{-/-}$ testes using the Fast RNA Pro Green Kit (MP Biomedicals) according to the manufacturer's protocol. Twelve-week-old animals were used unless stated otherwise. DNA remnants were removed using TURBO DNase (Ambion), and complementary DNA (cDNA) was made using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The quantitative PCRs were carried out on a StepOnePlus instrument (Applied Biosystems).

TaqMan® analysis

Reactions were set up using 50 ng cDNA, TaqMan® Fast Universal PCR Master Mix and appropriate TaqMan primers and probes (all reagents from Applied Biosystems). Default PCR program settings were used. Pre-designed primers and probes were applied for target genes (*Alkbh1*, *Tzfp*, *4933440M02Rik*) and endogenous control (*Gapdh*). See TaqMan probes 7–10 in Supplementary Table S1.

SYBR® green analysis

Reactions were set up using 5 ng cDNA, Power SYBR® Green PCR master mix and appropriate primers (300 nM) (all reagents from Applied Biosystems). Default PCR program and melt curve settings were used. Transposon primer sequences were similar to the ones used by Carmell *et al.* (7). Pachytene purity from STAPUT was verified using quantification of the pachytene marker *Lcn2* on isolated large RNAs (>200 nt). SYBR Green analysis with appropriate primers designed in http://eu.idtdna.com/SCITOOLS/Applications/PrimerQuest/Default.aspx were used. *GADPH* was used as endogenous control. See primers 17–28 in Supplementary Table S1.

At least two biological parallels were used for each genotype. All samples were run in triplicates and with two technical parallels. The relative expression was calculated using the equation $RQ = 2^{-\Delta\Delta CT}$, where RQ is the relative quantity of the target gene. $\Delta\Delta C_T$ is the difference in C_T values between target gene and the endogenous control minus the difference in C_T values between the reference gene and the endogenous control.

miScript PCR system for quantification of piRNAs

Small RNAs <200 nt were isolated from 2× C57BL6/J, *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ pachytene cells using the mirVana miRNA Isolation Kit (Ambion) in line with the manufacturer's protocol. Any DNA remnants were removed using TURBO DNase (Ambion), and cDNA was made using miScript Reverse Transcription Kit (Qiagen). The quantitative PCRs were carried out on a StepOnePlus instrument using 1 ng cDNA, miScript SYBR Green PCR Kit (Qiagen) and Custom miScript Primer Assays (piR-19852, piR-12359, piR-103121, piR-17918 and piR-4749) (Qiagen). The RNU6B miScript PCR control (snRNA, 45 nt) was chosen as endogenous control (Qiagen). All samples were run in triplicates with two technical parallels. The relative expression was calculated as for the TaqMan® Gene Expression Analysis. Primers 11–15 are listed in Supplementary Table S1.

Sequencing and computational analysis of small RNAs

Small RNAs from pachytene cells used in the miScript PCR system were also run in a high-throughput sequencing pipeline. Of RNA (<200 nt), 50 ng was diluted to 5.7 µl, entering the Small RNA Sample Preparation Guide (Illumina Part # 1004239 Rev. B August 2009) at the 5′-adapter ligation step. Subsequently, 3′-adapter ligation and cDNA synthesis followed by PCR amplification (15 cycles) were performed

using the Small RNA Sample Prep Kit (FC-102-1009, Illumina) according to the manufacturer's recommendations. Finally, the small RNA library (105–110 bp) was sequenced (36-bp Single-End Read) using the Illumina Genome Analyzer IIx (GAIIx).

Prior to analyzing, the sequence reads consisting solely of adapter sequences were removed. The remaining reads were aligned to the reference genome (mm9) using Novoalign (http://www.novocraft.com/), with settings allowing maximum one mismatch and automatic adapter trimming. For reads that had a best hit to fewer than six positions in the genome, all positions were considered. The rest were considered to have an unknown location. Known piRNA clusters were downloaded from piRNABank (31) and piRNAdb (http://kbrb.ioz.ac.cn/piRNA). All non-overlapping clusters from these two sources were considered for annotation. Other non-coding RNAs (miRNA, small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), transfer RNA (tRNA), long intergenic non-coding RNA (lincRNA), ribosomal RNA (rRNA) and miscellaneous RNA miscRNA) were downloaded from Ensembl using the BioMart system (http://www.ensembl.org/biomart/). In addition, GtRNAdb (http://gtrnadb.ucsc.edu/) was used for annotated tRNAs. Annotated exons, introns, 5′-UTR and 3′-UTR, in addition to the RepeatMasker track, were downloaded from the UCSC table browser.

Annotation of the mapped reads to the various non-coding RNA classes, as well as the gene and repeat tracks, was done using the intersectBed script from the BedTools software suit (32). The intersections were initially done against the annotated piRNA clusters. The resulting uncharacterized reads were then mapped to miRNA, rRNA, tRNA, snoRNA, snRNA, lincRNA and finally, miscRNA. Similarly, all reads were annotated to the RepeatMasker classes. The reads were organized into LTR, LINE and SINE. These reads were considered repeat derived. Any read that mapped to a different repeat category was classified as other repeat.

Potential piRNAs included reads that overlapped with the annotated piRNA clusters. The expression of piRNAs was measured at the cluster level by summing the number of mapped reads for each cluster intersected with intersectBed. Reads that mapped to several positions were divided between them such that fractions were used instead of full counts. To avoid counting unique sequences resulting from sequencing errors as novel piRNAs, we required at least two reads for a novel piRNA sequence to be called.

In order to detect significantly differentially expressed clusters, the DESeq package for the statistical environment R was used (33). Significantly differentially expressed clusters were defined as all clusters below a 10% false discovery rate.

Search for potential piRNA targets

Potential piRNA targets in the July 2007 (NCBI37/mm9) genome assembly of the *Mus musculus* strain C57BL6/J were identified using BLAST version 2.2.26 (34) with options set to detect highly diverged sequences (-W 7 -F F -r 3 -q -2 -G 5 -E 5). SWIPE version 2.0.3 (35) was also used for this purpose. The complete 4227-bp sequence of cluster 1082B was used as the query. Sequences matching only in the 97-bp 5′-LINE1 region or the adenine/guanine-rich region located at 2180–2470 bp were ignored. Matching regions were inspected using the UCSC genome browser (36).

RESULTS

Increased expression of a piRNA precursor in *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ testes

Alkbh1 deficiency in mice leads to reduced survival rate in embryos and causes embryonic abnormalities, and surviving mice have a dramatic sex-ratio skewing (26). Although male mice homozygous for the mutant gene were fertile, the sex-ratio distortion may possibly be due to aberrant formation of male germ cells. The hypothesis that Alkbh1 has a role in germ cell formation was strengthened by the high *Alkbh1* expression in testis.

To gain more insight into the role of Alkbh1 in mouse spermatogenesis, we searched for Alkbh1-regulated genes in adult testes using microarray analysis. Using class comparison strategy, we identified 25 differentially expressed genes in *Alkbh1*$^{-/-}$ testes. Although most of the genes affected are likely indirect targets of Alkbh1, one of the genes, *4933440M02Rik*, caught our interest due to its tremendous 15-fold upregulation (Figure 1A). The RefSeq for *4933440M02Rik* is suppressed as no support for a protein encoded by this transcript has been found. Parallel sequence alignment of *4933440M02Rik* in ParAlign 5.0 resulted in hits on 30 different mouse piRNAs. An expanded search in piRNABank (http://pirnabank.ibab.ac.in) revealed a cluster of 225 individual piRNAs containing the *4933440M02Rik* transcript on chromosome 7 (Figure 1B). The piRNA cluster is defined with id 1082 and has a cluster length of 39.2 kb.

The tremendous upregulation of the *4933440M02Rik* transcript in the *Alkbh1*$^{-/-}$ testes indicates a possible role for Alkbh1 in the regulation of the piRNA cluster located within the *4933440M02Rik* gene. This hypothesis was strengthened by the identification of a binding site for Tzfp, a possible Alkbh1 interaction partner, within the gene. Tzfp binds a specific genomic sequence, the tbs TG TACAGTGT (30), located within 'Exon 3' of the *4933440M02Rik* gene (Figure 1B). Tzfp is a known transcription repressor with a remarkably high expression in testis (37,38).

The ALKBH1–TZFP interaction was identified by screening a testis library by yeast two-hybrid analysis with full-length human ALKBH1 as bait (Supplementary figure S1A). The protein–protein interaction was verified using a second reporter gene (HTX assay, Supplementary figure S1A) as well as plasmid control digest (Supplementary figure S1B) and was further verified by dot blot analysis (Supplementary figure S1C). Although we have not been able to verify the ALKBH1-TZFP interaction through Co-IP analysis, the results indicate that these two proteins may be binding partners *in vivo*.

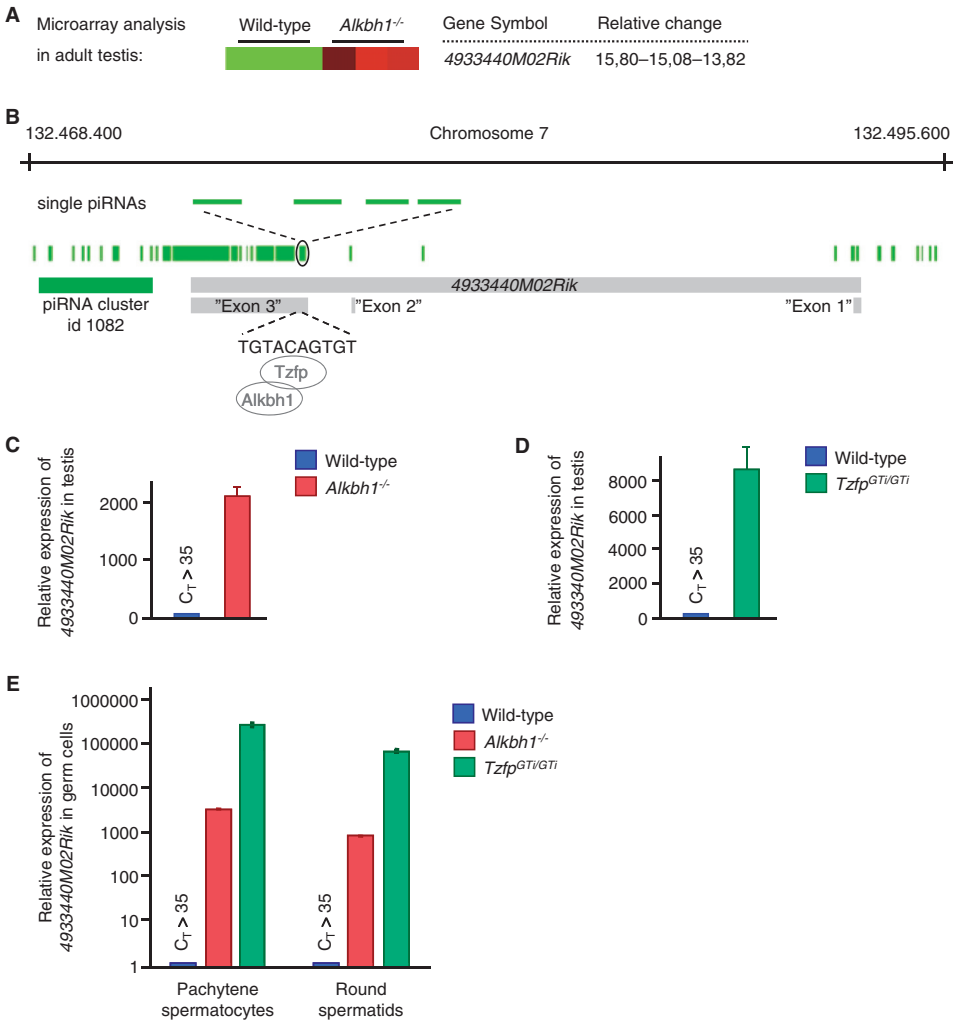


Figure 1. The expression of *4933440M02Rik* is upregulated in the $Alkbh1^{-/-}$ and $Tzfp^{GTi/GTi}$ testes. (**A**) Microarray analysis performed on whole testes indicating a 15-fold upregulation of the *4933440M02Rik* gene in $Alkbh1^{-/-}$ testes. (**B**) Schematic representation of the *4933440M02Rik* gene and transcript, a likely precursor for a piRNA cluster with id 1082 (adopted from http://pirnabank.ibab.ac.in). Magnification of indicated part of the cluster shows single piRNAs (light green) spliced from the piRNA precursor. The 10-bp Tzfp binding sequence (tbs) is located within 'Exon 3'. (**C**,**D**) qPCR analyses of the *4933440M02Rik* transcript in $Alkbh1^{-/-}$ (red) (C) and $Tzfp^{GTi/GTi}$ (green) (D) indicate a dramatic upregulation in the mutant testes when compared to wild type (blue). (**E**) qPCR analyses of the *4933440M02Rik* transcript in wild type (blue), $Tzfp^{GTi/GTi}$ (green) and $Alkbh1^{-/-}$ (red) pachytene spermatocytes (PC) and round spermatids (RS). Wild type is used as reference sample (RQ = 1).

The strong upregulation of the tbs-containing *4933440M02Rik* gene could imply that Tzfp, in co-operation with Alkbh1, represses the piRNA precursor (Figure 1B). To test this, we used the two mouse models $Alkbh1^{-/-}$ and $Tzfp^{GTi/GTi}$ to quantitate the *4933440M02Rik* transcript in $Alkbh1^{-/-}$ and $Tzfp^{GTi/GTi}$ testes. Indeed, quantitative PCR (qPCR) revealed a remarkable 2000-fold upregulation of the *4933440M02Rik* piRNA precursor in $Alkbh1^{-/-}$ testes (Figure 1C) and an 8000-fold upregulation in $Tzfp^{GTi/GTi}$ testes (Figure 1D). To determine whether this transcript is normally expressed during certain stages of spermatogenesis, male pups were killed at specific time points corresponding to the appearance of different spermatocytes and spermatids in the first wave of spermatogenesis in the juvenile mouse (day 10–26). In addition, pachytene spermatocytes and round

spermatids were isolated from adult male mice. qPCR analyses revealed low abundance of the *4933440M02Rik* transcript at all stages in juvenile mice (data not shown). In adult mutant cells, this transcript was detected at remarkably high levels in both pachytene spermatocytes and round spermatids, although a lower level was seen in round spermatids (Figure 1E, note the logarithmic scale on the *y*-axis). Since repression of the piRNA transcript requires both Alkbh1 and Tzfp, it seems likely that the regulation is caused by a stable or transient Alkbh1-Tzfp interaction at a specific DNA sequence.

Alkbh1 and Tzfp repress specific piRNAs in the pachytene stage of meiosis

It is previously shown that *Alkbh1* expression is particularly high in adult testis (26). For *Tzfp*, the testis-specific expression is especially pronounced (37,38). By analyzing RNA from juvenile testes, we find that both *Alkbh1* and *Tzfp* expression increases during pachynema (Figure 2A). The upregulation in the pachytene stage is more dramatic for *Tzfp* than for *Alkbh1*. Pachynema is the third stage of prophase I of meiosis, in which synapsis is completed and homologous recombination occurs.

Next, we wanted to investigate the expression profile of mature piRNAs derived from the area containing the *4933440M02Rik* gene during the pachytene stage. To address this question, we performed qPCR analysis on individual piRNAs (piRNA 1–5) from cluster 1082 on purified pachytene spermatocytes (Figure 2B). A dramatic increase in expression of piRNA 2–5 was found in both mutants (Figure 2C and D) and correlated very well with the expression of the *4933440M02Rik* gene described above (Figure 1E). Surprisingly, piRNA 1 expression was not affected by *Alkbh1* deletion and was slightly downregulated in the *Tzfp* mutant (Figure 2C and D). This piRNA is positioned furthest away from the tbs site and was found to be located outside the *4933440M02Rik* transcript, but within cluster 1082 as defined by piRNABank (http://pirnabank.ibab.ac.in).

Based on these results, Alkbh1 and Tzfp evidently regulate parts of cluster 1082 during pachynema of meiosis. The repression, however, does not comprise the whole length of this cluster. To get more detailed information on which area of cluster 1082 that is regulated, we performed high-throughput sequencing on small RNAs from $Alkbh1^{-/-}$, $Tzfp^{GTi/GTi}$ and wild-type pachytene cells. These studies will also uncover the possible regulation of other piRNA clusters by Alkbh1 and Tzfp.

The small RNA population is similar in wild-type and mutant pachytene spermatocytes

High-throughput sequencing of small RNAs isolated from purified pachytene spermatocytes was performed for the two mutants utilizing the 36-bp single-read sequencing on the Illumina Genome Analyzer IIx (GAIIx). Each mutant sample was analyzed along with a wild-type sample, resulting in the analysis of 2× six C57BL6/J adult mice and 1× six *Alkbh1* and *Tzfp* mutant mice. After adapter trimming and mapping of the sequences to the reference genome, the $Alkbh1^{-/-}$ sample yielded about 10 million reads, whereas the $Tzfp^{GTi/GTi}$ sample generated >15 million reads per sample (Supplementary Table S2). The number of reads is appreciably high and makes a promising basis for analysis with high statistical strength (13,39).

Small RNA reads were annotated as piRNAs if they mapped to a known piRNA cluster (see 'Materials and Methods' section). Our analysis demonstrated that pachytene piRNAs have a peak length of 29–30 nt (Figure 3A). Strikingly, as much as 93.7% of all sequence reads were piRNAs, whereas miRNAs only made up 1.4% of the matched reads (Figure 3B, left). The distribution of piRNAs and miRNAs was roughly the same in wild-type and $Tzfp^{GTi/GTi}$ pachytene spermatocytes (Figure 3B, right), whereas the miRNA population was slightly higher in $Alkbh1^{-/-}$ pachytene spermatocytes (Figure 3B, middle).

The sequencing data revealed a 5′-U bias and a variable 10th position in the pachytene piRNAs (data not shown), suggesting a lack of amplification by the 'ping-pong' mechanism. This is in accordance with previous findings and indicates that these piRNAs are generated through primary processing of precursor transcripts only (13,18).

Pachytene piRNAs can be divided into non-repeat, repeat-derived (LTR, LINE, SINE) and other repeat sequences (Figure 3C). Non-repeat piRNAs can be further subdivided into intergenic, exonic, intronic, 5′-UTR and 3′-UTR sequences (Figure 3D). We identified 82.7% of the wild-type piRNAs to be unique, i.e. non-repeat derived (Figure 3C) and 90.5% of the non-repeat piRNAs to be intergenic piRNAs (Figure 3D).

Based on the Illumina reads, we can conclude that the major portion of small RNAs in the meiotic pachytene stage is non-repeat intergenic piRNAs with a peak length of 29–30 nt. These findings are similar to other studies in adult testes from mouse and rat (9–11,13,14). In general, the profile of small RNAs was similar in wild-type and mutant pachytene spermatocytes, suggesting that Alkbh1 and Tzfp are not engaged in global piRNA regulation or biogenesis.

Cluster 1082B is the sole piRNA cluster regulated by Alkbh1 and Tzfp

Despite the strong upregulation seen for some of the piRNAs located within cluster 1082, deep sequencing of small RNAs from isolated pachytene spermatocytes revealed that the global small RNA profile was unchanged in mutant pachytene cells when compared to wild type. To investigate whether more piRNA clusters were differentially expressed, we first determined the expression profile of clusters by summing the number of mapped reads for each cluster intersected with intersectBed (see 'Materials and Methods' section). By mapping the sequence reads against known piRNA clusters from piRNABank (http://pirnabank.ibab.ac.in) and piRNAdb (http://kbrb.ioz.ac.cn/piRNA), a total of 4944 piRNA clusters were identified. Next, we established if any of the clusters were differentially expressed using the DESeq package for statistical testing (see 'Materials and Methods' section). Even though several piRNA clusters

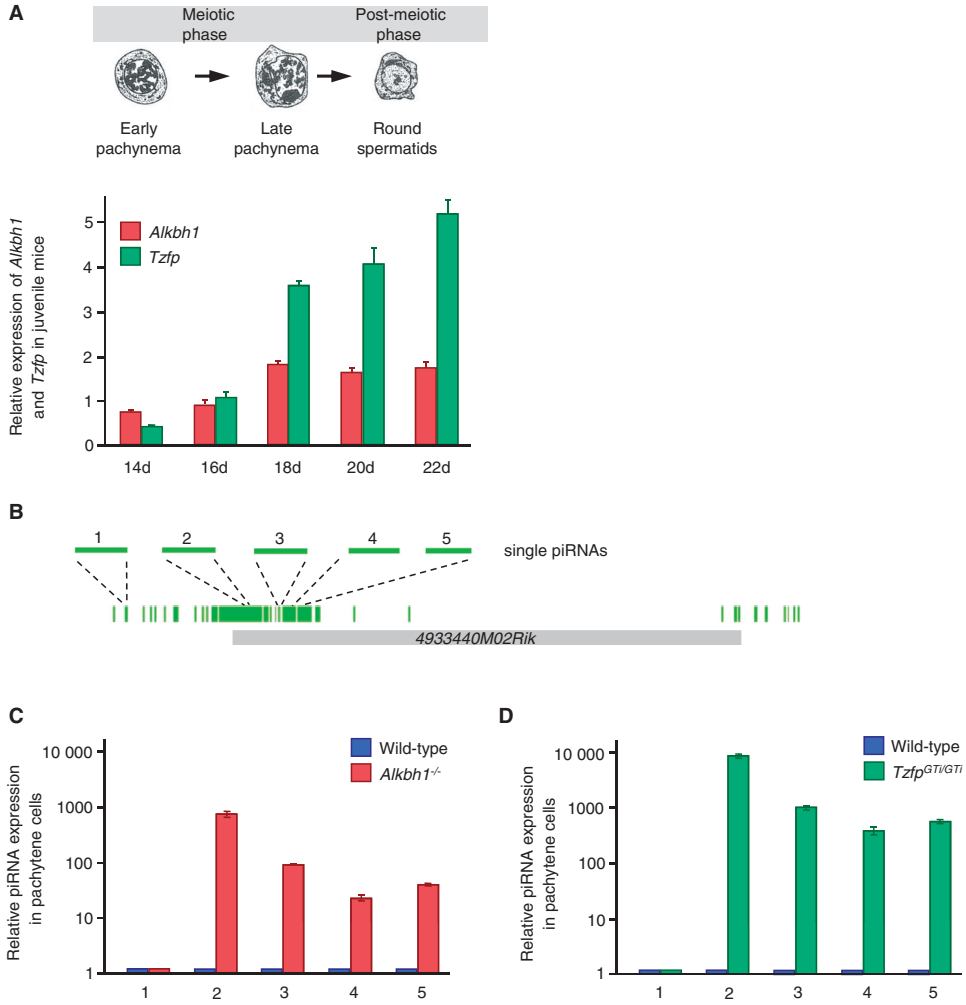


Figure 2. Individual piRNAs derived from parts of cluster 1082 are upregulated in *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ pachytene cells. (**A**) The upper panel shows a schematic drawing of the meiotic and post-meiotic phases during spermatogenesis (adopted from San Agustin and Witman, 2001). The appearance of early and late pachytene spermatocytes and round spermatids is indicated. The lower panel shows qPCR analyses of *Alkbh1* (red) and *Tzfp* (green) in male juvenile testes at 14–22 days after birth (d). Day 16, which is when mid-pachytene spermatocytes appear, is used as reference sample (RQ = 1). (**B**) Map showing the location of single piRNAs (light green) investigated by qPCR (numbered 1–5). The indicated piRNAs are distributed along the length of cluster 1082. (**C**, **D**) miScript qPCR analyses of the expression of piRNA 1–5 in isolated pachytene cells from wild type (blue), *Alkbh1*$^{-/-}$ (red) (C) and *Tzfp*$^{GTi/GTi}$ (green) (D). Wild type is used as reference sample (RQ = 1).

showed an increased or decreased expression in mutants compared to wild type (Supplementary Figure S3 and Supplementary Table S3), a part of cluster 1082 was the only region displaying a considerable upregulation in both *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ (Figure 4, Supplementary Table S3). This region was defined as cluster 1082B.

The sequence analysis showed that cluster 1082B covers a 4.2-kb sequence in chromosomal position 132 472 367–132 476 593 on chromosome 7 (Figure 5). Cluster 1082B is a relatively small piRNA cluster, both compared to cluster 1082, which originally had a cluster length of 39.2 kb, and to the average cluster length of 32 kb presented in a similar study (13). In our study, cluster 1082 is redefined into three separate clusters, called cluster 1082A, 1082B and 1082C (Figure 5). Cluster 1082A is a sense-strand cluster, whereas cluster 1082B and 1082C both consist of anti-sense piRNAs. This is in accordance with what Gan *et al.* found in their deep sequencing analyses (13).

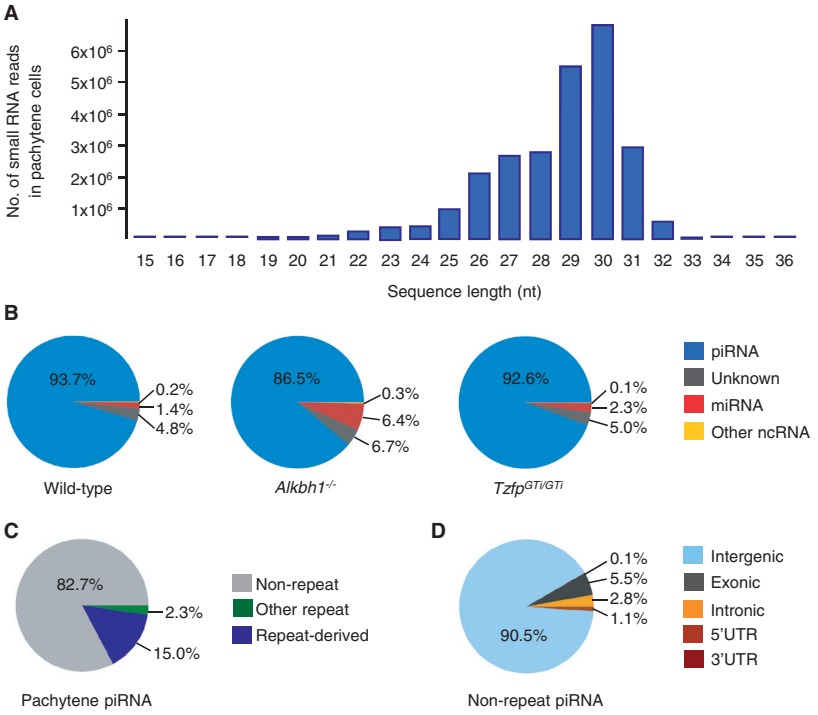


Figure 3. Small RNA sequencing analysis reveals a majority of 29–30 nt long non-repeat intergenic piRNAs in pachytene spermatocytes. (**A**) Sequence length distribution of small RNAs in wild-type pachytene cells after high-throughput sequencing, indicating peak expression of 29–30 nt long small RNAs. (**B**) Distribution of small RNA classes in isolated pachytene cells from wild type (left), *Alkbh1*$^{-/-}$ (middle) and *Tzfp*$^{GTi/GTi}$ (right). (**C**) Distribution of pachytene piRNAs into non-repeat associated, repeat associated (LTR, LINE and SINE) and other repeat sequences (analyses performed on the sample from (A)). (**D**) Distribution of non-repeat pachytene piRNAs into intergenic, exonic, intronic and UTR sequences (analyses performed on the sample from (A)).

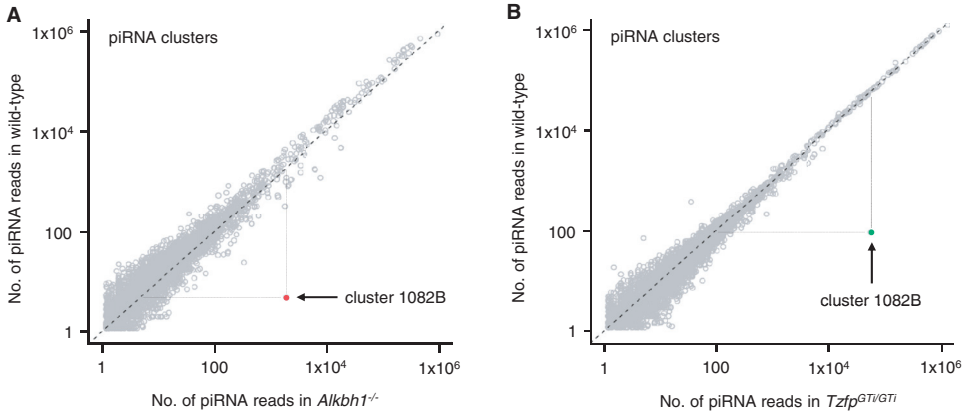


Figure 4. Cluster 1082B is the sole piRNA cluster upregulated in pachytene cells lacking Alkbh1 and Tzfp. Scatterplot analysis of piRNA cluster reads annotated by small RNA sequencing. Expression of piRNA clusters in pachytene cells is plotted on a log2 scale in a pairwise comparison of wild type (*y*-axis) and *Alkbh1*$^{-/-}$ (**A**) or *Tzfp*$^{GTi/GTi}$ (**B**) (*x*-axis). Clusters with a similar number of reads in the two samples line up along the center line (gray, open circles). Cluster 1082B, with a significantly higher number of reads in mutant samples than wild type, deviates noticeably from the center line (red, filled circle in (A); green, filled circle in (B)). The number of clusters presented is 4944. Dashed lines indicate the expression level in each genotype, showing an average expression level for cluster 1082B in wild type and a highly increased expression in the two mutants.

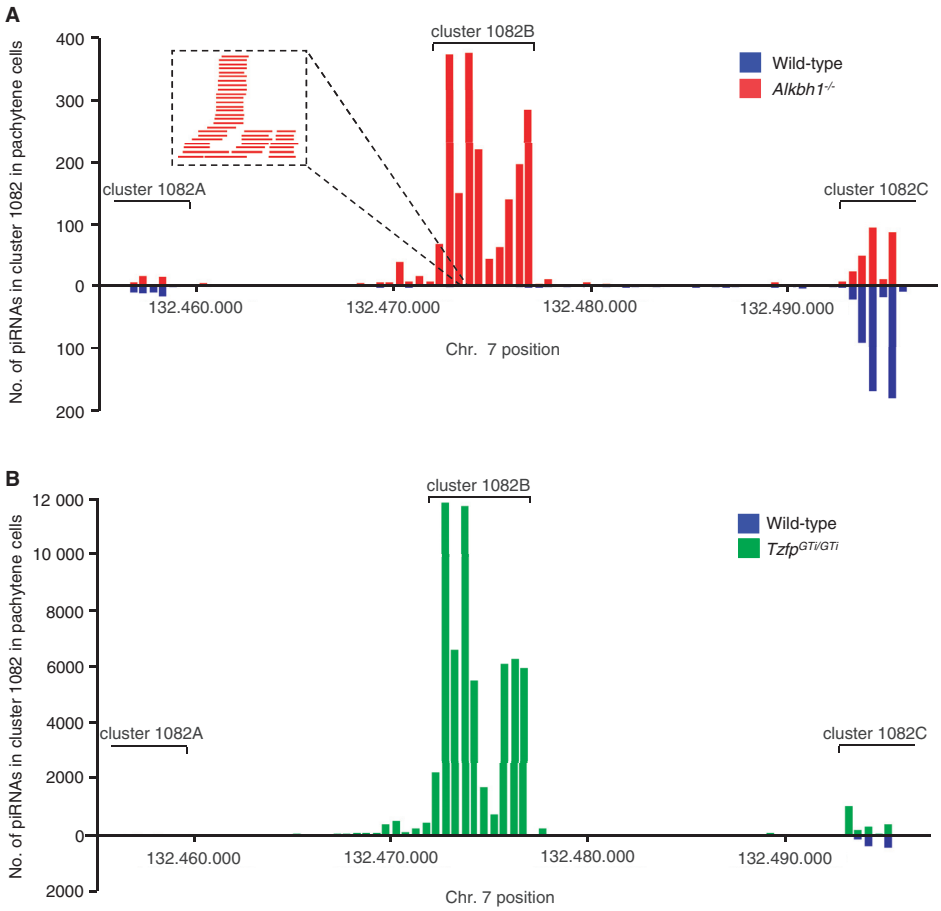


Figure 5. A 4.2-kb segment of piRNA cluster 1082 makes up cluster 1082B and is regulated by Alkbh1 and Tzfp. Chromosomal view of cluster 1082 on chromosome 7 with back-to-back histograms illustrating piRNA reads in wild type (blue bars) compared to *Alkbh1*$^{-/-}$ (red bars) (**A**) and *Tzfp*$^{GTi/GTi}$ (green bars) (**B**). After deep sequencing, cluster 1082 was redefined into the three separate clusters 1082A, 1082B and 1082C as indicated. Cluster 1082B is dramatically upregulated in mutant pachytene cells and corresponds to chromosomal position 132 472 367–132 476 593 containing 788 unique anti-sense piRNAs. A high-resolution view of a small segment of the upregulated region is given in (A) to illustrate individual, partially overlapping piRNAs.

The precursor sequence for cluster 1082B is partially located within the *4933440M02Rik* gene ('Exon 3') and contains the Tzfp recognition sequence. An 80% identical 1082B sequence is found on chromosome 1 in the rat genome, and the Tzfp recognition sequence is also conserved (data not shown). None of the other clusters investigated contained the tbs sequence (data not shown). The two mutants show an extreme upregulation in the expression of cluster 1082B, with a logfold change of 8.86 in *Alkbh1*$^{-/-}$ and 9.25 in *Tzfp*$^{GTi/GTi}$ compared to wild type (Figure 4 and Supplementary Table S3). Cluster 1082B contains 788 unique non-repeat piRNAs, and the distribution of the piRNAs along the cluster shows an almost identical distribution in *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ during pachynema (Figure 5). In wild-type pachytene spermatocytes, the expression of cluster 1082B is comparable to the expression level of other clusters investigated (Figure 4). In the mutant samples, however, cluster 1082B expression is highly above the expression level of most clusters. This means that many of the piRNA clusters are expressed at low levels and are only detectable with highly sensitive methods, e.g. deep sequencing.

We conclude that Alkbh1 and Tzfp specifically repress a 4.2-kb intergenic piRNA cluster on chromosome 7 during the pachytene stage of meiosis. The location of the tbs within this cluster indicates that the upregulated sequence represents a piRNA precursor transcript regulated by Alkbh1 and Tzfp.

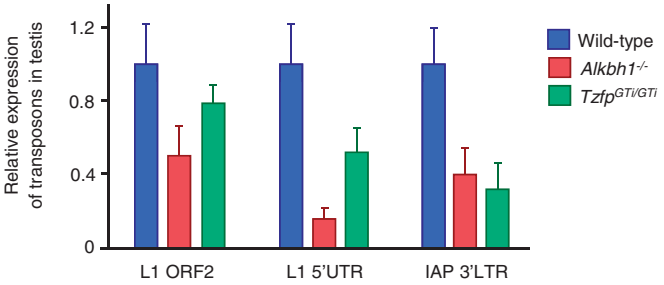


Figure 6. LINE1 and IAP transcripts are downregulated in the $Tzfp^{GTi/GTi}$ and $Alkbh1^{-/-}$ testis. qPCR analyses of transposable elements in adult testes from wild type, $Tzfp^{GTi/GTi}$ and $Alkbh1^{-/-}$ samples. The expressions of LINE1 (5′UTR and ORF2) and IAP (3′LTR) transcripts were tested. The two-tailed Student's *t*-test was used to check for significant difference in expression between wild type and mutants. A *P* value <0.001 was found for all three transcripts in both mouse models. Wild type is used as reference sample (RQ = 1).

Removal of Tzfp and Alkbh1 leads to LINE1 and IAP transcript downregulation

The detailed role of Miwi and its associated piRNAs have, until recently, been largely unknown, although a role in mRNA target silencing has been suggested (20). Reuter *et al*. recently showed that Miwi binds repeat-derived pachytene piRNAs in adult testes and cleaves target LINE1 RNA complementary to the piRNA sequence (18). This silencing was found to be ping-pong independent and not related to DNA methylation of LINE1 elements. The function of piRNAs derived from unannotated genomic regions, however, still remains a mystery. These piRNAs are not complementary to any other parts of the genome.

To investigate whether the 1082B-derived piRNAs might affect mRNA stability, we performed mRNA sequencing on wild-type and $Alkbh1^{-/-}$ pachytene spermatocytes. Even though five genes showed a more than 4-fold enrichment in the mutant preparation (Supplementary Figure S4 and Supplementary Table S4), none of the differentially expressed mRNAs are likely targets of the piRNAs from cluster 1082B. Next, we investigated whether LINE1 and IAP transcripts were affected in our two mouse models. Surprisingly, qPCR analysis of IAP (3′-LTR) and LINE1 (5′-UTR and ORF2) elements revealed significant downregulation in the $Tzfp^{GTi/GTi}$ and $Alkbh1^{-/-}$ testes, with an average reduction of ∼50% ($P < 0.001$) (Figure 6). This could indicate that the piRNAs regulated by Alkbh1 and Tzfp in pachytene spermatocytes have the potential to silence LINE1 and IAP sequences and that piRNA upregulation seen in the $Tzfp^{GTi/GTi}$ and $Alkbh1^{-/-}$ mice thus causes a reduction in transposon transcripts.

Search for potential piRNA targets

To further investigate the possible function of the piRNAs derived from cluster 1082B, a homology search was performed to see whether these piRNAs can bind to other transcripts or genomic regions. Similarity searches using the 4227-bp sequence of cluster 1082B as query identified several potential piRNA target regions in the mouse genome. Potential targets with significant sequence similarity were identified on chromosomes 3, 6, 11, 12, 13, 16, 17 and 18 (Figure 7 and Supplementary Table S5). The best match was to a region on chromosome 13 that includes the gene *4921525O09Rik* ($E = 10^{-45}$). Interestingly, Girard *et al*. and Lau *et al*. have previously mapped several piRNAs to *4921525O09Rik* (9,10), indicating that it probably contains a piRNA precursor transcript. The regions on chromosome 11 and 17 include the pseudogene *4932414J04Rik* and the gene *gm324*, respectively, both encoding putative serine/threonine protein kinases. The region on chromosome 16 includes a gene (*4933404G15Rik*) with unknown function.

DISCUSSION

The high abundance of piRNAs in immature male germ cells indicates that they have a vital biological role. To study the possible regulation of piRNAs by Alkbh1 and Tzfp, we performed small RNA sequencing on pachytene spermatocytes isolated from mutant and wild-type testes. We obtained a notable ∼15 million reads and ∼10 million reads per sample after adapter trimming and mapping of the sequences to the reference genome, allowing high-quality analysis.

The mouse Piwi protein Miwi is detected from the mid-pachytene stage to elongating spermatids (8). Mili is expressed from primordial germ cells until the round spermatid stage (8). piRNAs expressed during the pachytene stage of meiosis can thus bind both Miwi and Mili, resulting in at least two subpopulations of piRNAs in adult testes (8,9,20). IP analysis has shown that Mili-interacting piRNAs peak at 26–28 nt, whereas Miwi bind 29–31 nt long piRNAs (8). Using deep sequencing analyses, we could confirm that pachytene piRNAs have a peak length of 29–30 nt, indicating that they primarily interact with Miwi (40). In accordance with other studies, we found the majority of small RNAs in pachytene spermatocytes to be piRNAs derived from non-repeat associated, intergenic sequences (13).

ALKBH1 localizes to nuclear euchromatin and recent evidence shows that ALKBH1 is a dioxygenase targeting a novel methylated lysine in histone H2A (27)

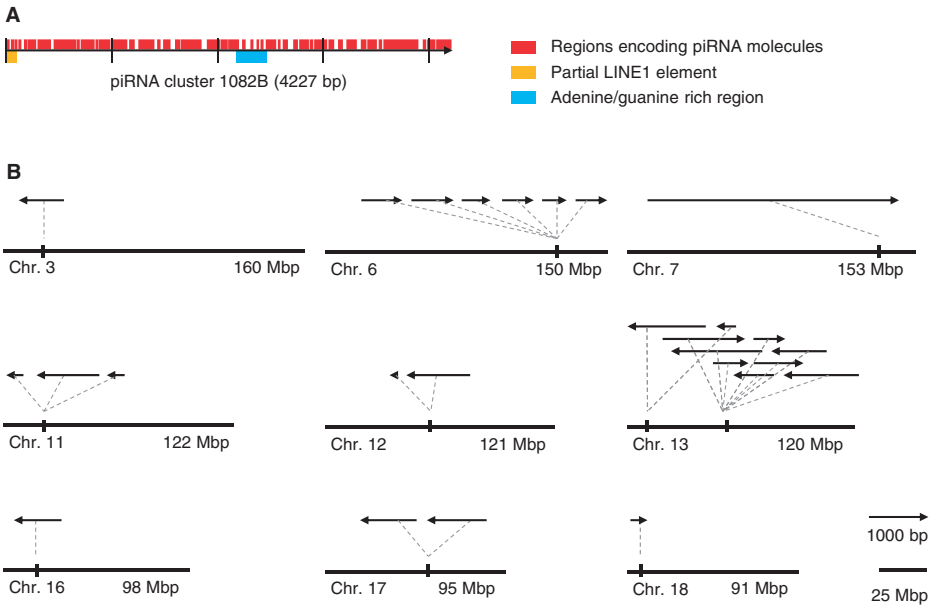


Figure 7. Potential target regions in the mouse genome of piRNAs from cluster 1082B identified by sequence similarity. (**A**) Location of regions encoding piRNAs (red), a partial LINE1 element (1–97 bp) (yellow) and an adenine/guanine-rich region (2180–2470 bp) (blue) in the sequence of cluster 1082B. (**B**) The chromosomal location, direction and length of potential piRNA target regions identified by homology search using the 1082B sequence as query. Chromosome 7 with cluster 1082B is included for reference. Distances between target regions are not drawn to scale.

(Ougland *et al*., submitted for publication), suggesting a role for the protein in gene regulation through an epigenetic mechanism (26–28). Tzfp, a member of the C2H2-type zinc finger family, binds to the tbs, leading to repression of target genes (30). Here, we show that both proteins are highly expressed during pachynema in spermatogenesis. Using deep sequencing analyses, we found that piRNAs derived from a non-repeat cluster on chromosome 7, called cluster 1082B, are highly upregulated in *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ pachytene spermatocytes. The non-coding RNA transcript covering cluster 1082B was also greatly upregulated in mutant testes, indicating a role for Tzfp and Alkbh1 in the regulation of the precursor transcript encoding the piRNAs within this region. This hypothesis is further substantiated by the identification of the 10-bp Tzfp recognition sequence contained within the cluster. No other piRNA cluster was found to be clearly differentially expressed, indicating that Tzfp and Alkbh1 are not involved in global piRNA regulation.

Combined, these results indicate that Alkbh1 and Tzfp together regulate cluster 1082 in testis. We believe that the repression of the piRNA precursor occurs by Tzfp binding to the tbs sequence, while Alkbh1 changes the surrounding chromatin environment by histone H2A demethylation. During the progress of this project, we also generated the *Alkbh1*$^{-/-}$/*Tzfp*$^{GTi/GTi}$ double knockout mice. These mice had reduced fertility and viability comparable to the single *Alkbh1*$^{-/-}$ knockout mouse (data not shown). This suggests an epistatic relationship between *Alkbh1* and *Tzfp*. The regulatory mechanism of pachytene piRNA clusters has previously been unknown, but our results indicate that piRNA precursor transcripts are individually regulated by stage-specific proteins determining the timing and amount of piRNA expression. Gan *et al*. found that a large number of transcripts expressed during spermatogenesis, including pseudogenes, are piRNA precursors and suggest that these genes are regulated by alternative splicing and anti-sense transcripts (13). Our findings invite us to speculate whether some of the long ESTs with unknown functions in mouse testes are single-stranded piRNA precursors regulated by members of the C2H2-type zinc finger transcription factors.

Despite intense research over the past few years, the precise function of pachytene piRNAs is still largely unknown. As these piRNAs originate from clusters spread throughout the genome, one of the functions may be patronage of their respective loci, as seen for scnRNAs in *Tetrahymena* (41), but other functions such as post-transcriptional regulation and roles in epigenetic programming cannot be ruled out. It is now well established that Mili and Miwi2 are essential for *de novo* DNA methylation and repression of LINE1 in fetal and newborn mice (7,14,16), but as only ~17% of adult piRNAs map to annotated repeats in mammals (11,14), it has been unclear whether this occurs in adult testes. Interestingly, Grivna *et al*. observed that Miwi is associated with polysomes throughout spermatogenesis and that it associates with several mRNAs through the

cap binding complex (11,20). This would suggest that Miwi-associated piRNAs may be capable of silencing target mRNA transcripts. Other studies have shown that some piRNAs are located in the meiotic nucleus of the male mouse, indicating that these piRNAs could be involved in chromosome remodeling during meiosis and post-meiotic stages (42,43).

The present study can give new insight into the mechanisms related to pachytene piRNAs, as our results indicate that pachytene piRNAs derived from non-repeat sequences may have a role in transposon control. This hypothesis is based on the downregulation of LINE1 and IAP transcripts seen in the *Alkbh1*$^{-/-}$ and *Tzfp*$^{GTi/GTi}$ testes. Although it cannot be ruled out that the downregulation is a secondary effect, it is tempting to speculate that these piRNAs target LINE1 and IAP transcripts, leading to reduced expression. Reuter *et al.* recently showed that Miwi and its associated repeat-derived piRNAs target LINE1 RNAs in adult testes, resulting in transcript cleavage by the endonuclease activity of Miwi (18). The cleavage was found to be dependent upon high degree of complementarity between piRNA and target sequence. Yet, the 1082B-derived piRNAs, like the majority of pachytene piRNAs, are devoid of sequences complementary to active transposons, indicating a slicer-independent mechanism. Further studies are needed to elucidate the mechanism by which these piRNAs work, but a miRNA-like mechanism is possible. Upon miRNA binding to target, only a small part of the 5′-end, called the miRNA 'seed', requires Watson–Crick pairing (44,45). Binding of the miRNA to the target transcript leads to repression rather than direct degradation.

As silencing of transposons is regarded important for the maintenance of a cell's integrity, it is somewhat surprising that a cluster of piRNAs, with the potential to inactivate transposons, is repressed in wild-type pachytene cells. LINE1 is, however, known to be expressed in wild-type germ cells (46), and it is possible that a certain level of LINE1 expression during spermatogenesis is evolutionary beneficial. It has been speculated that LINE1 insertions may structurally modify endogenous genes and regulate gene expression as they move and duplicate over evolutionary time (47).

A homology search performed to find putative targets of the piRNAs derived from cluster 1082B yielded highly significant hits in several regions of the genome. None of the regions included LINE1 or IAP sequences and did not overlap with any of the piRNA clusters identified in our deep sequencing analysis. The region with highest degree of homology overlapped with the gene *4921525O09Rik*, which has previously been shown to contain several piRNAs (9,10). Some of these piRNAs are anti-sense to the cluster 1082B primary transcript, indicating that *4921525O09Rik*- and 1082B-derived piRNAs could bind to sense and anti-sense transcripts in target regions. So far, the biological relevance of non-repeat piRNAs is unclear, and more research is needed to identify the mechanism through which they work. Studies have revealed that the genomes of mammals are almost entirely transcribed (48,49), giving rise to a wide range of intergenic, anti-sense, overlapping and intronic non-protein-coding RNAs. Some of these transcripts could be regulatory RNAs, taking part in the networked interactions between transcription initiation and elongation, modulation of chromatin architecture, RNA editing, alternative splicing, RNA and protein signaling. Our findings emphasize the complexities of gene and chromatin regulation during differentiation in mammals.

In summary, we identified a piRNA cluster located in a non-repeat intergenic region of chromosome 7 that encodes a unidirectional piRNA precursor transcript giving rise to 788 unique anti-sense piRNAs. The cluster is transcriptionally repressed in wild type but was found to be the only cluster highly upregulated in *Tzfp*$^{GTi/GTi}$ and *Alkbh1*$^{-/-}$ pachytene spermatocytes, indicating a role for Tzfp and Alkbh1 in the regulation of this cluster. The dramatic upregulation of the 1082B-derived piRNAs was followed by a downregulation of LINE1 and IAP transcripts in testes from both mutants. It is thus tempting to speculate that the piRNAs derived from cluster 1082B are involved in transposon control during the pachytene stage of meiosis in male germ cells.

ACCESSION NUMBERS

GSE37150.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–5, Supplementary Figures 1–4, Supplementary Methods and Supplementary Reference [50].

AVAILABILITY

The raw sequence data from this study have been submitted to the NCBI Gene Expression Omnibus under accession number GSE37150.

ACKNOWLEDGEMENTS

We are grateful to Merete Worren from the Bioinformatics Core Facility at Oslo University Hospital for helping out with the mRNA sequencing analysis. We also thank Linda Ellevog for excellent technical assistance with the germ cell isolations and qPCR analysis. We thank the Centre for Comparative Medicine at Oslo University Hospital for the excellent service they provided.

FUNDING

Funding for open access charge: The Norwegian Research Council and the Norwegian Cancer Society.

Conflict of interest statement. None declared.

REFERENCES

1. Hess,R.A. and Renato de,F.L. (2008) Spermatogenesis and cycle of the seminiferous epithelium. *Adv. Exp. Med. Biol.*, **636**, 1–15.
2. Russell,L.B. (1990) Patterns of mutational sensitivity to chemicals in poststem-cell stages of mouse spermatogenesis. *Prog. Clin. Biol. Res.*, **340C**, 101–113.
3. Cox,D.N., Chao,A., Baker,J., Chang,L., Qiao,D. and Lin,H. (1998) A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.*, **12**, 3715–3727.
4. Lin,H. and Spradling,A.C. (1997) A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development*, **124**, 2463–2476.
5. Deng,W. and Lin,H. (2002) miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell*, **2**, 819–830.
6. Kuramochi-Miyagawa,S., Kimura,T., Ijiri,T.W., Isobe,T., Asada,N., Fujita,Y., Ikawa,M., Iwai,N., Okabe,M., Deng,W. *et al.* (2004) Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development*, **131**, 839–849.
7. Carmell,M.A., Girard,A., van de Kant,H.J., Bourc'his,D., Bestor,T.H., de Rooij,D.G. and Hannon,G.J. (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell*, **12**, 503–514.
8. Aravin,A., Gaidatzis,D., Pfeffer,S., Lagos-Quintana,M., Landgraf,P., Iovino,N., Morris,P., Brownstein,M.J., Kuramochi-Miyagawa,S., Nakano,T. *et al.* (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, **442**, 203–207.
9. Girard,A., Sachidanandam,R., Hannon,G.J. and Carmell,M.A. (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, **442**, 199–202.
10. Lau,N.C., Seto,A.G., Kim,J., Kuramochi-Miyagawa,S., Nakano,T., Bartel,D.P. and Kingston,R.E. (2006) Characterization of the piRNA complex from rat testes. *Science*, **313**, 363–367.
11. Grivna,S.T., Beyret,E., Wang,Z. and Lin,H. (2006) A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.*, **20**, 1709–1714.
12. Thomson,T. and Lin,H. (2009) The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu. Rev. Cell Dev. Biol.*, **25**, 355–376.
13. Gan,H., Lin,X., Zhang,Z., Zhang,W., Liao,S., Wang,L. and Han,C. (2011) piRNA profiling during specific stages of mouse spermatogenesis. *RNA*, **17**, 1191–1203.
14. Aravin,A.A., Sachidanandam,R., Girard,A., Fejes-Toth,K. and Hannon,G.J. (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*, **316**, 744–747.
15. Aravin,A.A., Sachidanandam,R., Bourc'his,D., Schaefer,C., Pezic,D., Toth,K.F., Bestor,T. and Hannon,G.J. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell*, **31**, 785–799.
16. Kuramochi-Miyagawa,S., Watanabe,T., Gotoh,K., Totoki,Y., Toyoda,A., Ikawa,M., Asada,N., Kojima,K., Yamaguchi,Y., Ijiri,T.W. *et al.* (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.*, **22**, 908–917.
17. Klattenhoff,C. and Theurkauf,W. (2008) Biogenesis and germline functions of piRNAs. *Development*, **135**, 3–9.
18. Reuter,M., Berninger,P., Chuma,S., Shah,H., Hosokawa,M., Funaya,C., Antony,C., Sachidanandam,R. and Pillai,R.S. (2011) Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature*, **480**, 264–267.
19. De,F.S., Bartonicek,N., Di,G.M., breu-Goodger,C., Sankar,A., Funaya,C., Antony,C., Moreira,P.N., Enright,A.J. and O'Carroll,D. (2011) The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature*, **480**, 259–263.
20. Grivna,S.T., Pyhtila,B. and Lin,H. (2006) MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. *Proc. Natl Acad. Sci. USA*, **103**, 13415–13420.
21. Unhavaithaya,Y., Hao,Y., Beyret,E., Yin,H., Kuramochi-Miyagawa,S., Nakano,T. and Lin,H. (2009) MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. *J. Biol. Chem.*, **284**, 6507–6519.
22. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089–1103.
23. Gunawardane,L.S., Saito,K., Nishida,K.M., Miyoshi,K., Kawamura,Y., Nagami,T., Siomi,H. and Siomi,M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. *Science*, **315**, 1587–1590.
24. Ragan,C., Mowry,B.J. and Bauer,D.C. (2012) Hybridization-based reconstruction of small non-coding RNA transcripts from deep sequencing data. *Nucleic Acids Res*, June 16 (doi:10.1093/nar/gks505; epub ahead of print).
25. Siomi,M.C., Sato,K., Pezic,D. and Aravin,A.A. (2011) PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.*, **12**, 246–258.
26. Nordstrand,L.M., Svard,J., Larsen,E., Nilsen,A., Ougland,R., Furu,K., Lien,G.F., Rognes,T., Namekawa,S.H., Lee,J.T. *et al.* (2010) Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects. *PLoS One*, **5**, e13827.
27. Pan,Z., Sikandar,S., Witherspoon,M., Dizon,D., Nguyen,T., Benirschke,K., Wiley,C., Vrana,P. and Lipkin,S.M. (2008) Impaired placental trophoblast lineage differentiation in Alkbh1(−/−) mice. *Dev. Dyn.*, **237**, 316–327.
28. Kurowski,M.A., Bhagwat,A.S., Papaj,G. and Bujnicki,J.M. (2003) Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. *BMC Genomics*, **4**, 48.
29. Trewick,S.C., Henshaw,T.F., Hausinger,R.P., Lindahl,T. and Sedgwick,B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature*, **419**, 174–178.
30. Tang,C.J., Lin,C.Y. and Tang,T.K. (2006) Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. *Dev. Biol.*, **290**, 398–410.
31. Sai,L.S. and Agrawal,S. (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res.*, **36**, D173–D177.
32. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.*, **26**, 841–842.
33. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
34. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
35. Rognes,T. (2011) Faster Smith-Waterman database searches with inter-sequence SIMD parallelisation. *BMC Bioinformatics*, **12**, 221.
36. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
37. Lin,W., Lai,C.H., Tang,C.J., Huang,C.J. and Tang,T.K. (1999) Identification and gene structure of a novel human PLZF-related transcription factor gene, TZFP. *Biochem. Biophys. Res. Commun.*, **264**, 789–795.
38. Hoatlin,M.E., Zhi,Y., Ball,H., Silvey,K., Melnick,A., Stone,S., Arai,S., Hawe,N., Owen,G., Zelent,A. *et al.* (1999) A novel BTB/POZ transcriptional repressor protein interacts with the Fanconi anemia group C protein and PLZF. *Blood*, **94**, 3737–3747.
39. Ma,L., Buchold,G.M., Greenbaum,M.P., Roy,A., Burns,K.H., Zhu,H., Han,D.Y., Harris,R.A., Coarfa,C., Gunaratne,P.H. *et al.* (2009) GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *PLoS. Genet.*, **5**, e1000635.
40. Aravin,A.A., Hannon,G.J. and Brennecke,J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, **318**, 761–764.
41. Mochizuki,K. and Gorovsky,M.A. (2004) Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan

(scn) RNAs involved in genome rearrangement. *Genes Dev.*, **18**, 2068–2073.
42. Marcon,E., Babak,T., Chua,G., Hughes,T. and Moens,P.B. (2008) miRNA and piRNA localization in the male mammalian meiotic nucleus. *Chromosome Res.*, **16**, 243–260.
43. Beyret,E. and Lin,H. (2011) Pinpointing the expression of piRNAs and function of the PIWI protein subfamily during spermatogenesis in the mouse. *Dev. Biol.*, **355**, 215–226.
44. Lewis,B.P., Burge,C.B. and Bartel,D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
45. Brennecke,J., Stark,A., Russell,R.B. and Cohen,S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.
46. Branciforte,D. and Martin,S.L. (1994) Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell Biol.*, **14**, 2584–2592.
47. Han,J.S. and Boeke,J.D. (2005) LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? *Bioessays*, **27**, 775–784.
48. Cheng,J., Kapranov,P., Drenkow,J., Dike,S., Brubaker,S., Patel,S., Long,J., Stern,D., Tammana,H., Helt,G. *et al.* (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*, **308**, 1149–1154.
49. Bertone,P., Stolc,V., Royce,T.E., Rozowsky,J.S., Urban,A.E., Zhu,X., Rinn,J.L., Tongprasit,W., Samanta,M., Weissman,S. *et al.* (2004) Global identification of human transcribed sequences with genome tiling arrays. *Science*, **306**, 2242–2246.
50. Trapnell,C., Pachter,L. and Salzberg,S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105–1111.

III

Tzfp represses the Androgen Receptor in mouse testis

Kari Furu[1] and Arne Klungland[1,2,#]

1 Centre for Molecular Biology and Neuroscience, Department of Microbiology, Oslo University Hospital, Rikshospitalet, PO Box 4950 Nydalen, NO-0424 Oslo, Norway, **2** Institute of Basic Medical Sciences, University of Oslo, PO Box 1018 Blindern, NO-0315 Oslo, Norway

Corresponding author[#]: Arne Klungland

E-mail: arne.klungland@medisin.uio.no

Keywords: Tzfp, pachytene spermatocytes, Androgen Receptor

Abstract

The testis zinc finger protein (Tzfp), also known as Repressor of GATA, belongs to the BTB/POZ zinc finger family of transcription factors and is thought to play a role in spermatogenesis due to its remarkably high expression in testis. Despite many attempts to find the *in vivo* role of the protein, the molecular function is still largely unknown. Here, we address this issue using a novel mouse model with a disrupted *Tzfp* gene. Mice lacking Tzfp are viable and fertile, but display an increase in Androgen Receptor (AR) signaling in Sertoli cells and altered expression of several genes in the testis. Our results indicate that Tzfp represses *Gata1*, *Aie1* and *Fancc*, leading to altered AR signaling, resulting in a reduced number of apoptotic cells in the testicular tubules.

Introduction

The Testis Zinc Finger Protein (Tzfp), also known as Repressor of GATA, is a transcription factor belonging to the BTB/POZ-ZF (Broad complex, Tramtrack, Brick à brack (BTB) or poxvirus and zinc finger (POZ)-zinc finger) protein family [1,2]. The BTB domain is a widely distributed protein-protein interaction motif that is often found at the N-terminus of zinc finger transcription factors, whereas the C terminal Krüppel C_2H_2 zinc fingers are thought to be an important DNA-binding domain. The BTB/POZ-ZF proteins generally function as transcription repressors and are involved in a broad range of biological processes, including cell survival and differentiation [3,4]. Tzfp is homologous to another BTB/POZ-ZF protein, Plzf (promyelocytic leukemia zinc finger protein). BLAST analysis reveals over 50% identity, with the highest similarity in the

zinc finger domain. PLZF is regarded as a tumor suppressor and works as a transcriptional repressor and chromatin remodeller implicated in cancer and the self-renewal of spermatogonial stem cells [5,6].

Mice deficient of Tzfp have previously been generated [7,8], revealing that the protein is not embryonic lethal and that mice lacking Tzfp develop without any serious phenotype. In blood, TZFP is expressed at high levels in the early stages of differentiation but declines during subsequent differentiation into erythroid and myeloid lineages [9]. Disruption of the *Tzfp* gene leads to increased T lymphocyte proliferation, cytokine production and altered hematopoetic stem cell homeostasis [8], whereas overexpression of the protein is accompanied by accumulation in G1 and increased rates of apoptosis. Recent findings also suggest a role in B lymphocyte differentiation [10]. TZFP makes a heterodimer with the FA group C protein (FANCC) [11], which is one of at least 13 distinct complementation proteins involved in the autosomal recessive disorder Fanconi anemia (FA) [12]. Tzfp also binds GATA-3, one of six members in the GATA transcription factor family [13]. GATA-3 is involved in the development of the T cell lineage and TZFP represses GATA-3-induced transactivation. These results suggest a role for Tzfp during the proliferative stages of primitive hematopoetic progenitors; however, its precise roles in other forms of cell differentiation remain unknown. The protein is, however, thought to play a role in spermatogenesis due to its high expression in testis [1,11,13].

Tzfp probably exists in at least two isoforms [1,13,14]. Isoform 1 consists of 465 amino acids (Fig. 1A) and is predicted to be testis specific, whereas the shorter isoform 2 is expressed in B-and T lymphoid cells and lacks the C-terminal zinc finger motif. In 2001,

Tang *et al.* showed that the zinc finger domain of isoform 1 binds to a specific genetic sequence, the Tzfp binding site (tbs) TGTACAGTGT, in the upstream flanking sequence of the *Aie1* gene [15]. They also showed that the BTB domain has a repressive effect, indicating that Tzfp negatively regulates the expression of genes carrying the tbs sequence. *Aie1* encodes Aurora-C, a member of the Aurora kinase protein family, and is primarily expressed in germ cells where it is thought to regulate chromosome segregation during male meiosis [16].

A link between TZFP and the Androgen Receptor (AR) was recently discovered by Kaufmann *et al.* who show that TZFP can form a trimeric complex with AR and the endogenous retrovirus encoded protein Rec [17]. AR is a member of the nuclear receptor superfamily and functions as a ligand-dependent transcription factor, regulating the expression of an array of androgen-responsive genes [18,19]. AR and androgens are vital for spermatogenesis [20], and can be detected in Sertoli cells (SC), the Peritubular myoid (PM) cells, and cells in the interstitial spaces including Leydig cells and perivascular smooth muscle cells in testis [21,22].

In the present study we use Tzfp deficient mice to investigate the role of this protein in spermatogenesis. In agreement with previous studies, we find a remarkably high expression of *Tzfp* in testis [1,11,13]. We further show that *Tzfp* is most highly expressed in pachytene spermatocytes during prophase I of meiosis. Mice lacking the *Tzfp* gene were viable, fertile, and exhibited no obvious phenotypic defects. Moreover, we show that removal of Tzfp leads to altered regulation of *Fancc, Aie1*, and *GATA-1* in testis, and that it affects Androgen Receptor signaling in Sertoli cells. These results show that Tzfp

is a potent repressor of several genes in murine testis and that it is a part of a complex system that ensures proper regulation of meiosis.

Materials and methods

Generation of the *Tzfp*$^{GTi/GTi}$ mouse. Embryonic stem (ES) cells containing a gene trap cassette inserted upstream of the *Tzfp* exon 1 were obtained from the Texas Institute of Genomic Medicine (C57Bl/6 gene trap ES cell clone IST12443F7). Briefly, the gene trap vector includes a promoter-less reporter gene (β-GEO) downstream of a splice acceptor (SA) sequence followed by a polyA tail. Insertion of the gene trap cassette terminates transcription prematurely, resulting in a transcript containing none of the exons within the *Tzfp* gene. We injected the IST12443F7 ES cells into C57Bl/6 blastocysts and obtained male mice capable of transmitting the mutated *Tzfp* allele.

The gene trap insertion site was determined using iPCR-based direct-sequencing protocol. ES cells were microinjected into blastocysts and implanted into recipient female mice to permit development of the embryos into chimeras, and subsequent breeding was performed to select founders with germ line transmission of the gene trap allele. The mice were maintained on a C57Bl/6 background.

All mouse experiments were approved by the Norwegian Animal Research Authority (Ref. nr. 08/109551) and done in accordance with institutional guidelines at the Centre for Comparative Medicine at Oslo University Hospital. Animal work was conducted in accordance with the rules and regulations of the Federation of European Laboratory Animal Science Association's (FELASA). Mice were weaned between 3 and 4 weeks of

age, and genotyped by PCR on genomic DNA prepared from ear clips. Genomic DNA was prepared by incubation in 75 μL Hot Shot Lysis Buffer (25 mM NaOH, 0.2 mM Na_2EDTA, pH 12) at 95°C for 30 min, and then cooled down to 4°C before adding 75 μL Hot Shot Neutralization Buffer (40 mM Tris-HCl, pH 5). Samples were PCR amplified for 40 cycles with an annealing temperature of 58°C using the following primers:

Wild-type allele Tzfp gene:	5'-ACTGTGGCAGACTAATACTT-3'
	5'-GCTCAACAAGTCAAGACTTT-3'
Targeted allele Tzfp gene:	5'-ACTGTGGCAGACTAATACTT-3'
	5'-CTTGCAAAATGGCGTTACTTAAGC-3'

StaPut Isolation of Pachytene Cells. Pachytene spermatocytes and round spermatids were isolated from 12-week old testes using an adapted version of the StaPut method [23]. Six 12-week old male pups were sacrificed using CO_2 and the testes were detunicated and the tubules treated with DNase1 (200μg) and collagenase (10mg) in 10ml DMEM for 8 min at 34°C. The tubules were then incubated with DNase1 (200μg), collagenase (10mg), trypsin (10mg) and hyaluronidase (15mg) in 10 ml DMEM for 15 min at 34°C (all enzymes from Sigma-Aldrich). These steps remove connective tissue and extratubular cells, yielding a cell suspension of germinal cells in DMEM containing 0.5% BSA. The cell suspension was loaded into the cell loading chamber of the StaPut apparatus and subsequently separated by sedimentation velocity at unit gravity in a 2-4% w/v BSA gradient in DMEM medium at 4°C for 2.5 hours. After sedimentation, 10 ml fractions were collected, and the fractions containing pachytene spermatocytes and round spermatids were found based on size and morphological features. Pure fractions were pooled and the cell number counted on a Countess® Automated Cell Counter

(Invitrogen). The cells were spun down and the cell pellet was snap frozen in liquid nitrogen. An aliquot of purified cells was fixed on glass slides using Cell Adherence Solution (Crystalgen) on SuperFrost Plus slides (VWR) for microscopic analyses. One separation yielded approximately 1.5×10^6 pachytene cells with an average diameter of 12.5 µm and 1×10^7 round spermatids with an average diameter of 9.9 µm.

TaqMan® Gene Expression analysis. Total RNA was isolated from cells and tissue using the Fast RNA Pro Green Kit (MP Biomedicals) or the mirVana™ miRNA Isolation Kit (Ambion) according to the manufacturers' protocol. Any DNA remnants were removed using TURBO DNase (Ambion), and cDNA syntesis was made using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The RT-PCR reactions were carried out on a StepOnePlus instrument using 50 ng cDNA, TaqMan® Fast Universal PCR Master Mix and appropriate TaqMan primers and probes (all from Applied Biosystems). GAPDH was used as endogenous control in all assays except in the mouse tissue analysis where 18s was used. The following pre-designed TaqMan® probes were used:

GAPDH: Mm99999915_g1
18s: Hs99999901_s1
Tzfp: Mm00491292_g1
Fancc: Mm01301576_m1
AuroraC: Mm03039428_g1
Mll4: Mm01175395_g1
Plzf: Mm01176865_m1
Gata-1: Mm01352636_m1
Gata-4: Mm00484689_m1

Tissue and cells from 12 week old animals were used unless stated otherwise. The expression profiling was done using the relative quantitative method $\Delta\Delta C_T$. All samples were run in triplicates and with two technical parallels (2 runs per sample). A minimum of two biological parallels were used for each organ, and testis from 6 animals were pooled when performing analysis on germ cells. Statistical analysis was performed using non-paired, two-tailed Student's *t* test.

***In Situ* Hybridzation.** Testes from 12 week old mice were removed and subsequently placed in Tissue-tek O.C.T Compound (Sakura) on dry ice. The samples were cut into 15 µm thick slices onto SuperFrost Plus slides (VWR) and allowed to air dry. Sections were labeled with DIG-RNA sense and anti-sense probes, covering an area of 801 bases (base 238 to 1038) of the *Tzfp* transcript. The probes were made using a PCR product with T7 and T3 sequence flanking the two strands, and DIG RNA labelling mix (Roche). The following primers were used:

Forward primer with T3 sequence: 5'-**aattaaccctcactaaagg**ggagagagtatagagctaca-3'
Reverse primer with T7 sequence: 5'-**taatacgactcactatagg**gcgagcaggagagggtaaag-3'

In Situ hybridization was performed using an adapted version of the protocol by Hoover and Goldman [24]. The sections were fixed in 4% paraformaldehyde, dehydrated in 100-50% ethanol and rinsed in 2×SSC (150mM NaCl and 15mM sodium citrate, pH 7.4). Permeabilization was performed using 10 µg/ml Proteinase K (Roche) in 0.1M Tris–HCl, pH 7.5 with 50mM EDTA for 15 min at 37°C. Sections were post-fixed in 4% paraformaldehyde and treated with 0.2M HCl for before acetylation in 0.1M triethanolamine (TEA), pH 7.5, with 0.25% acetic anhydride. Sections were dehydrated

and air dried before incubation with hybridization solution (10mM Tris–HCl, pH 7.5, 50% formamide, 0.3M NaCl, 1mM EDTA, 10% dextran sulfate (Fluka) and 1% blocking solution (Roche)), containing 100 ng DIG-RNA probe. Slides were incubated at 55°C over night and then rinsed. Unhybridized RNA was digested by incubation with RNase (50µg/ml) in NTE buffer (10mM Tris pH 7.5, 0.5M NaCl, 1mM EDTA) for 30 min at 37°C. Slides were then washed in the NTE buffer for 30 min at 60°C. Unspecific staining was blocked by incubating in 2×SSC with 0.05% Triton X-100 and 1% blocking solution for 2 hours before incubation with anti-DIG-alkaline phosphate (1:1000) (Roche) in MAB buffer (100mM maleic acid, 150mM NaCl, pH 7.5) over night at 4°C. The sections were washed in MAB and incubated with equilibration buffer (100mM Tris–HCl, 100mM NaCl and 50mM MgCl) before detection with NBT (Nitro blue tetrazolium chloride) and BCIP (5-bromo-4-chloro-3-indolyl phosphate, toluidine salt) tablets (Roche), and counterstained with haematoxylin. Pictures were taken using an AxioCam ICc1 camera on an Axio Observer.Z1 microscope (Carl Zeiss).

Immunofluorescent staining of frozen testis sections. Ten µm thick frozen sections were cut onto SuperFrost Plus slides (VWR) and allowed to air dry. Before staining, the sections were fixed in 4% Pfa for 10 minutes, permeabilized in PBS with 0.5% Triton for 5 minutes and blocked in PBS containing 5% BSA and 5% normal Goat Serum (Invitrogen) for 30 min at room temperature. The slides were then incubated with primary antibodies overnight at 4°C or 37°C prior to detection with secondary antibodies. Primary antibodies used were rabbit anti-AR (1:50, Abcam), rabbit anti-GATA-1 (1:100, Cell Signaling), chicken anti-laminin (1:50, Abcam). Secondary antibodies used were donkey anti-rat Alexa Fluor 594 (1:500, Invitrogen), and goat anti-rabbit Alexa 594 (1:500,

Invitrogen). The slides were counterstained with DAPI (Invitrogen) and mounted with Mowiol (Merck Biosciences Ltd). Pictures were taken using an AxioCam MR Rev3 camera on an Axio Observer.Z1 microscope with Apotome (Carl Zeiss). The images were processed using AxioVision 4.8 software (Carl Zeiss) and Image J.

Signal intensity measurements. Signal intensity upon immunohistochemistry was measured in selected cells using the ImageJ software. The same conditions and exposure times were used on all samples when comparing cells on different slides. Cells of interest were defined using the Region of Interest (ROI) applications, and a mean intensity was found measuring several cells from three different tubules from each genotype. Statistical analysis was performed using non-paired, two-tailed Student's *t* test.

Haematoxylin and eosin staining. Paraffin embedded mouse testes were cut in 4 μm thick slices on a microtome and fixed on SuperFrost Plus slides (VWR). The slides were deparaffinated in ClearRite3 and rehydrated in 100-70% ethanol. Staining was performed by incubation in haematoxylin 7211 for 2 min, Blueing Reagent 7301 for 1 min and eosin 71204E for 1.5 min (all solutions from Richard-Allan Scientific). Slides were dehydrated, washed in ClearRite 3 and mounted using Mounting Medium 4111 (Richard-Allan Scientific). Pictures were taken using an AxioCam ICc1 camera on an Axio Observer.Z1 microscope (Carl Zeiss) and the images were processed using AxioVision 4.8 software (Carl Zeiss) and Image J.

Apoptosis detection. Paraffin embedded tissue sections from wild-type and knock-out mouse testis were made and deparaffinated and rehydrated as described above. The sections were quickly rinsed in PBS before Proteinase K treatment (20μg/ml) for 20

minutes at 37°C. TUNEL assay was performed according to the manufacturer's protocol using Cell Death detection kit, TMR red (Roche). The sections were counterstained with DAPI and mounted using Mowiol (Merck Biosciences Ltd). Slides from at least two different animals from each genotype were used. Pictures were taken using an AxioCam MR Rev3 camera on a microscope Axio Observer.Z1 with Apotome (Carl Zeiss). Slides from 3 different animals from each genotype were used and >100 tubules from each genotype were evaluated and the number of positive cells determined. The average number of apoptotic cells was then determined. Statistical analysis was performed using non-paired, two-tailed Student's *t* test.

Western Blot. Proteins were isolated from wild-type and *Tzfp*$^{GTi/GTi}$ testis using RIPA buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 0.5% NP-40, 0.1% SDS, 0.5 mM PMSF (AppliChem) and 1x Protease Inhibitor Coctail (Sigma)). The protein concentration was measured using Protein assay dye reagent concentrate (BioRad) and 150 µg protein was loaded onto NuPAGE® 10% Bis-Tris Gel 1.0 mm (Invitrogen). After SDS-PAGE, the proteins were blotted onto a PVDF membrane and the membrane was blocked in PBS with 5% skim milk before incubating with primary antibody over night at 4°C. A HRP-conjugated secondary antibody was used and the proteins detected using Immune Star WesternC kit (BioRad).

Primary antibodies used are: Anti-AR (Abcam, 1:100), anti-GATA-1 (Cell Signaling, 1:1000), anti-GATA-4 (Santa Cruz, 1:100) and anti-β-actin (Abcam, 1:2000). Secondary antibodies used are: HRP-conjugated anti-rabbit (Abcam, 1:20,000) and HRP-conjugated anti-goat (Abcam, 1:20,000).

Results

Tzfp deficient mice are viable and fertile. A Tzfp mutant mouse was generated using C57BL/6 ES cells containing a splice acceptor site from the gene trap vector (Omnibank Gene Trap Vector 76) upstream of exon 1 of the *Tzfp* gene (FIG 1A). Genotyping confirmed the presence of heterozygous animals in the first litter and later homozygous animals (FIG 1B). RT-PCR analysis showed an average CT value reduction of 9 in the homozygous mutant mouse compared to wild-type (FIG 2B). This indicates a decrease in transcript amount of >99%. Mice homozygous for the insertion are hereafter called $Tzfp^{GTi/GTi}$. Absence of the *Tzfp* transcript was also confirmed through *In Situ* hybridization on testis sections (FIG 2E).

We identified another gene located just 1.0 kb upstream from the *Tzfp* gene. This neighboring gene, *Mll4* (mixed-lineage leukemia), is suggested to function as a methyltransferase capable of methylating H3K4 [25]. The expression of *Mll4* was analyzed by RT-PCR to make sure that the expression of this gene was not affected in the $Tzfp^{GTi/GTi}$ mouse. As described above, Tzfp is homologous to Plzf, a protein encoded by the *Zfp145* gene. The expression of this gene was analyzed to confirm that Plzf does not work as a backup protein in our model, masking any effect of the *Tzfp* deletion. We saw no significant changes in *Mll4* or *Zfp145* expression in our mutant ($p>0.05$) (FIG 1C).

We intercrossed heterozygous animals and obtained normal, fertile homozygous pups with no apparent phenotype. The observation that Tzfp is not essential for the survival of mice pups or the formation of gametes, is in concurrence with the findings of other groups [7,8]. During the first generations of heterozygous breeding (number of

litters=10), we observed a deviation from the expected frequencies of sex- and genotype distributions (FIG 1D). Transmission ratio distortion (TRD) led to a frequency of homozygous null mice of only 10 % rather than the expected 25% ($p=0.02$ using χ^2 test). In addition, the average litter size was 5.1, significantly smaller than the average litter size of 7.0 pups per litter for C57BL/6J mice, indicating loss of homozygous embryos. We also saw a sex ratio distortion yielding 70% female pups rather than the expected 50% ($p=0.003$ using χ^2 test). The male homozygous null mice had testes of size and weight slightly lower than average ($p<0.05$), but normal epididymus ($p>0.5$) (FIG 1E).

The Tzfp transcript is most highly abundant in pachytene spermatocytes in the testis. We found *Tzfp* to be most highly expressed in testes (FIG 2A). This is in agreement with the findings of other groups [1,11,13]. To find out which cells in the male mouse germline expressing this gene, we isolated pachytene spermatocytes and round spermatids using unit gravity sedimentation from adult animals (12 weeks). RT-PCR analysis showed *Tzfp* expression to be highly elevated in pachytene spermatocytes (FIG 2B), where the gene is upregulated by approximately 50 fold compared to whole testis. The transcript can also be detected in round spermatids, although at a lower level. There is no expression in the vas deferens and epididymus, indicating absence of *Tzfp* in mature spermatozoa.

To further investigate the Tzfp expression, C57BL/6J male pups were sacrificed at specific time points corresponding to the appearance of different spermatocytes and spermatids in the first wave of spermatogenesis in the juvenile mouse [23,23]. The *Tzfp* expression rapidly increases as spermatogenesis progresses through prophase I (FIG 2C). The *Tzfp* transcript amount increases rapidly from day 16, indicating high expression

from mid-pachynema [26,23]. The expression continues to increase as more cells expressing *Tzfp* develop, and peaks at day 24. To verify the high expression of *Tzfp* in pachytene spermatocytes, we performed *In Situ* hybridization on frozen sections using a DIG-labeled RNA probe (FIG 2D). Most tubules have a clear ring of positive cells, indicating *Tzfp*-expressing pachytene cells. There is no signal in the $Tzfp^{GTi/GTi}$ samples and in the samples incubated with sense-probe.

Removal of the Tzfp protein leads to alterations in gene expression. To identify a possible pathway in which Tzfp is involved, RNA was isolated from juvenile (12 dpp when *Tzfp* expression is low/absent) and adult testes (12 weeks when *Tzfp* expression is high), and the gene expression pattern of several candidate genes was analyzed. A possible link between Tzfp and the Androgen Receptor has been found [17], yet no significant difference was found in the expression pattern of the AR transcript between wild-type and $Tzfp^{GTi/GTi}$ testes (FIG 3). TZFP has also been linked to the Fanconi Anemia pathway due to its ability to bind to the FA factor FANCC. Our RT-PCR analysis show that *Fancc* is significantly upregulated in the adult $Tzfp^{GTi/GTi}$ testis compared to the wild-type ($p<0.001$) (FIG 3).

We also wanted to investigate whether loss of *Tzfp* has an effect on the *Aie1* expression. As described above, Tang *et al.* have proposed a mechanism in which Tzfp has a represses this gene by binding to the tbs motif in the *Aie1* flanking region [15]. Our analyses showed a significant upregulation of *Aie1* expression in $Tzfp^{GTi/GTi}$ adult testes when compared to wild-type ($p<0.001$) (FIG 3), confirming the repressive effect on this gene by Tzfp.

Tzfp has been described as a repressor of GATA [13], and we therefore studied the expression levels of *Gata* mRNAs and proteins in $Tzfp^{GTi/GTi}$ testis. GATA-1, 4 and 6 are expressed in the Sertoli cells of the testis where they serve to drive spermatogenesis [27,28,29]. GATA-1 is expressed in a stage-specific manner, where the strongest signal is found in the Sertoli cells in stage VII-XI seminiferous tubules [30]. RT-PCR analysis revealed that *Gata1* expression is significantly upregulated ($p<0.001$) in adult $Tzfp^{GTi/GTi}$ testis when compared to wild-type (FIG 3), indicating a role for Tzfp in the regulation of GATA-1 in testes. A slight downregulation was found in the expression of *Gata-4* in adult testis ($p=0.01$) (FIG 3). As expected, no differences in gene expression were observed in juvenile testes where Tzfp is absent (FIG 3).

Removal of the Tzfp protein leads to increased amount of Androgen Receptor signaling in Sertoli cells. The Androgen Receptor is important for normal progression through spermatogenesis and is expressed by the Sertoli cells and the Peritubular myoid cells (PTMs) within the testicular tubules [31,32,33]. As described above, no significant differences in *AR* expression were found between wild-type and $Tzfp^{GTi/GTi}$ testes (FIG 3), indicating that Tzfp has no effect on *AR* transcription. However, when staining wild-type and $Tzfp^{GTi/GTi}$ testis sections with anti-AR, a stronger signal in the nuclei of Sertoli cells in the $Tzfp^{GTi/GTi}$ samples were detected (FIG 4A). This indicates that AR localization, turnover and/or activity are altered upon removal of Tzfp. This, in turn, may indicate a possible role for Tzfp as an AR regulator. As expected, anti-AR revealed similar AR levels in wild-type and *Tzfp* deficient juvenile testes. Quantification of the signal intensity in Sertoli cells from adult testis revealed a significant difference ($p=0.001$) between $Tzfp^{GTi/GTi}$ and wild-type samples (FIG 4C). The data was confirmed by Western analysis,

showing AR upregulation in the *Tzfp*$^{GTi/GTi}$ samples (FIG 4D). The strongest AR signal was seen in stage VII-VIII tubuli, which contain preleptotene and mid-pachytene spermatocytes, and step 7-8 and 16 spermatids. A strong AR signal in these stages has previously been established [22], indicating a crucial role in AR signaling in these cell types.

Despite the upregulation of *Gata-1* observed in *Tzfp*$^{GTi/GTi}$ testes, immunostaining with anti-GATA-1 revealed no difference in signal intensity when comparing *Tzfp*$^{GTi/GTi}$ testis sections with wild-type (FIG 4B-C). No significant difference was detected for GATA-1 and GATA-4 with Western analysis (FIG 4D), indicating that lack of Tzfp does not alter the production of these GATA proteins in the testis.

Tzfp null mice display normal progression through spermatogenesis, but have a decreased amount of tubular apoptosis. We investigated wild-type and *Tzfp*$^{GTi/GTi}$ testis by staining sections with haematoxylin and eosin (HE) to see whether the increased AR signal and changed gene expression in the *Tzfp*$^{GTi/GTi}$ mouse testis results in abnormalities in testis anatomy. Stage VII-VIII tubules were of particular interest as these stages contain cells with high *Tzfp* expression, are androgen sensitive, and have high expression of GATA-1. The histological sections, however, looked normal and appear to contain all spermatogenic cell types (FIG 5A). This is not surprising, as the mice are fertile and display normal sexual behavior. The tubular and luminal area was measured and we found that both parameters were slightly smaller in the *Tzfp*$^{GTi/GTi}$ testis compared to the wild-type ($p= 0.01$ and 0.001, respectively) (data not shown). This is in accordance with the slightly smaller testis size found in the *Tzfp*$^{GTi/GTi}$ mouse (FIG 1E). We also counted

the number of Sertoli cells per tubuli and the Sertoli cell nuclear area in wild-type and adult *Tzfp*$^{GTi/GTi}$ testis. No differences were found (data not shown).

We wanted to see whether the *Tzfp*$^{GTi/GTi}$ testis had an altered rate of apoptosis in the testis tubules. For this purpose we performed a TUNEL assay followed by statistical testing to see if absence of Tzfp gave increased cell death. Surprisingly, the number of positive cells was lower in *Tzfp*$^{GTi/GTi}$ mice compared to in wild-type (p=0.02) (FIG 5B).

Discussion

In the present study we show that removal of the Tzfp protein is not lethal, but that breeding mice heterozygous and homozygous for the gene trap insertion yields fewer *Tzfp*$^{GTi/GTi}$ pups than expected, indicating decreased survival of *Tzfp* deficient embryos. Although the protein has its highest expression in testis, it is dispensable for the formation of germ cells in both the male and female mouse. The gene is most highly expressed in pachynema, which is the stage in prophase I of meiosis where the synapsis of homologous chromosomes is completed and crossing over occurs.

It is assumed that Tzfp acts as a transcriptional repressor, as most BTB/POZ-ZF proteins are transcription factors with a repressive effect on their target genes. Several BTB/POZ-ZF proteins work as transcriptional regulators through the activation of HDACs via recruitment of the transcriptional co-repressors SMRT and N-CoR [34,35]. Transcriptional repression is a crucial part of the cell- and phase specific gene regulation necessary for normal meiosis.

We found *Fancc*, a gene encoding the FA factor FANCC, to be upregulated in $Tzfp^{GTi/GTi}$ mouse. Mice mutant for components in the Fanconi Anemia pathway show reduced fertility, probably due to failure in the mitotic proliferation in primordial germ cells [36,37]. FANCC functions in homologous recombination, and promotes the mutational repair of endogenously generated abasic sites [38]. One might speculate that upregulation of the FA pathway disturbs commitment of damaged cells to undergo apoptosis. Indeed, a reduced number of apoptotic cells was observed in $Tzfp^{GTi/GTi}$ testes. Analogous, FANCC prevents apoptosis in hematopoietic cells [39].

$Tzfp^{GTi/GTi}$ testes have increased Androgen Receptor signaling in the nuclei of Sertoli cells in testicular tubules when compared to wild-type. The signal was found to be increased 3-fold, indicating that Tzfp has a direct or indirect repressive effect on AR. Sertoli cells are mitotic cells located within the seminiferous tubules of the testis. These cells branch out ensuring intimate contact with the developing sperm, supporting and nurturing the cells. Sertoli cells are important for all phases of gametogenesis, including germ cell proliferation, meiosis and differentiation [40]. Studies with a Sertoli cell selective AR knockdown (SCARKO) have revealed that the presence of AR in Sertoli cells is vital for a successful completion of spermatogenesis [41,42]. In the SCARKO mouse, no germ cells develop to elongated spermatids, rendering mutant male mice infertile. This is probably due to the strong dependence on androgens for mid-pachytene spermatocytes and stage 7 spermatids, which are found in stage VII-VIII tubuli [22,32]. AR signaling fluctuates with the spermatogenic cycle and peaks in these stages [22].

The GATA factors are other candidates for proteins under Tzfp control that may regulate AR activity. The gonads are prominent sites of GATA expression, with detectable levels

of Gata-1, 2, 4 and 6 [43,44]. Of these factors only GATA-2 is expressed in the germ-cells [43], whereas Sertoli cells express GATA-1, 4 and 6. As described above, Tzfp is known to be a repressor of GATA, and an interaction between TZFP and the GATA factors 2 and 3 have been shown [13,45]. However, all vertebrate GATA proteins share a conserved DNA and protein-binding domain composed of two zinc fingers. This leads to an apparent redundancy in binding properties to target sequences [46], indicating that other GATA members might also bind TZFP.

GATA-1 exerts a temporal expression in Sertoli cells [44,28] and is only found in stage VI–XII tubules of the seminiferous cycle [30,47]. This phenomenon appears to be dependent on the presence of maturing germ cells. The expression of GATA-1 coincides with the androgen-sensitive stages of the seminiferous cycle [22] and with the stages containing cells with a high *Tzfp* expression. A 3-fold upregulation of *Gata-1* was apparent in $Tzfp^{GTi/GTi}$ testis, indicating that Tzfp has a repressive effect on the *Gata-1* gene. This makes GATA-1 perhaps the most likely candidate for a protein under Tzfp control to regulate AR activity. The mechanism through which this occurs is currently unknown, but it is possible that the repression occurs by the recruitment of HDACs. GATA-1 interact with several HDAC proteins [48], of which HDAC1 has been shown to bind AR and specifically down-regulate AR transcriptional activity [49]. GATA-1 also regulates the androgen receptor corepressor 19kDa (ARR19), which co-translocates into the nucleus with AR and suppresses AR activity through the recruitment of HDAC4 [50,51].

Regulation of AR by Tzfp could also occur via Aurora-C, a member of the Aurora kinase protein family. An interaction between Tzfp and the Aurora-C-encoding gene, *Aie1*, has

previously been established [15] and in addition, a significant upregulation of this gene was apparent in the *Tzfp*$^{GTi/GTi}$ testis when compared to wild-type. The Aurora kinase proteins are serine/threonine kinases essential for cell proliferation and are believed to play important roles in chromosome segregation [52]. Aurora-A interacts with and phosphorylates AR in prostate cancer cells, leading to increased AR activation [53]. The specificity of Aurora-C remains to be identified. Recently, it was reported that overexpression of Aurora-C in cell lines induces abnormal cell division resulting in centrosome amplification and multinucleation [54]. Such morphological abnormalities were not observed in the *Tzfp*$^{GTi/GTi}$ samples, indicating that the *Aie1* overexpression is not dramatic enough to cause problems with meiotic divisions.

In summary, we generated a viable and fertile *Tzfp*$^{GTi/GTi}$ mouse using a gene trap ES clone. We found the gene to be dispensable for life and the formation of germ cells and that the gene is most highly expressed during pachynema, a stage in meiotic prophase I characterized by the occurrence of crossing over. Tzfp most probably has a regulatory role during spermatogenesis and most likely functions as a transcriptional repressor. Removal of the protein leads to a significant increase in the expression of the genes *Fancc, Aurora-C* and *Gata-1,* and an increased AR signaling in Sertoli cells.

Reference List

1. Lin W, Lai, CH, Tang, CJ, Huang, CJ, and Tang, TK (1999) Identification and gene structure of a novel human PLZF-related transcription factor gene, TZFP. Biochem.Biophys.Res.Commun. 264: 789-795.

2. Perez-Torrado R, Yamada, D, and Defossez, PA (2006) Born to bind: the BTB protein-protein interaction domain. Bioessays 28: 1194-1202.

3. Melnick A, Ahmad, KF, Arai, S, Polinger, A, Ball, H et al. (2000) In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. Mol.Cell Biol. 20: 6550-6567.

4. Kelly KF and Daniel, JM (2006) POZ for effect--POZ-ZF transcription factors in cancer and development. Trends Cell Biol. 16: 578-587.

5. Barna M, Merghoub, T, Costoya, JA, Ruggero, D, Branford, M et al. (2002) Plzf mediates transcriptional repression of HoxD gene expression through chromatin remodeling. Dev.Cell 3: 499-510.

6. Costoya JA, Hobbs, RM, Barna, M, Cattoretti, G, Manova, K et al. (2004) Essential role of Plzf in maintenance of spermatogonial stem cells. Nat.Genet. 36: 653-659.

7. Kang BY, Miaw, SC, and Ho, IC (2005) ROG negatively regulates T-cell activation but is dispensable for Th-cell differentiation. Mol.Cell Biol. 25: 554-562.

8. Piazza F, Costoya, JA, Merghoub, T, Hobbs, RM, and Pandolfi, PP (2004) Disruption of PLZP in mice leads to increased T-lymphocyte proliferation, cytokine production, and altered hematopoietic stem cell homeostasis. Molecular and Cellular Biology 24: 10456-10469.

9. Dai MS, Chevallier, N, Stone, S, Heinrich, MC, McConnell, M et al. (2002) The effects of the Fanconi anemia zinc finger (FAZF) on cell cycle, apoptosis, and proliferation are differentiation stage-specific. J.Biol.Chem. 277: 26327-26334.

10. Yoon HS, Scharer, CD, Majumder, P, Davis, CW, Butler, R et al. (2012) ZBTB32 Is an Early Repressor of the CIITA and MHC Class II Gene Expression during B Cell Differentiation to Plasma Cells. J.Immunol.

11. Hoatlin ME, Zhi, Y, Ball, H, Silvey, K, Melnick, A et al. (1999) A novel BTB/POZ transcriptional repressor protein interacts with the Fanconi anemia group C protein and PLZF. Blood 94: 3737-3747.

12. Moldovan GL and D'Andrea, AD (2009) How the fanconi anemia pathway guards the genome. Annu.Rev.Genet. 43: 223-249.

13. Miaw SC, Choi, A, Yu, E, Kishikawa, H, and Ho, IC (2000) ROG, repressor of GATA, regulates the expression of cytokine genes. Immunity. 12: 323-333.

14. Ikeda R, Yoshida, K, and Inoue, I (2007) Identification of FAZF as a novel BMP2-induced transcription factor during osteoblastic differentiation. J.Cell Biochem. 101: 147-154.

15. Tang CJ, Chuang, CK, Hu, HM, and Tang, TK (2001) The zinc finger domain of Tzfp binds to the tbs motif located at the upstream flanking region of the Aie1 (aurora-C) kinase gene. J.Biol.Chem. 276: 19631-19639.

16. Tang CJ, Lin, CY, and Tang, TK (2006) Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. Dev.Biol. 290: 398-410.

17. Kaufmann S, Sauter, M, Schmitt, M, Baumert, B, Best, B et al. (2010) Human endogenous retrovirus protein Rec interacts with the testicular zinc-finger protein and androgen receptor. J.Gen.Virol. 91: 1494-1502.

18. Eder IE, Haag, P, Basik, M, Mousses, S, Bektic, J et al. (2003) Gene expression changes following androgen receptor elimination in LNCaP prostate cancer cells. Mol.Carcinog. 37: 181-191.

19. Jiang F and Wang, Z (2003) Identification of androgen-responsive genes in the rat ventral prostate by complementary deoxyribonucleic acid subtraction and microarray. Endocrinology 144: 1257-1265.

20. Maclean JA and Wilkinson, MF (2005) Gene regulation in spermatogenesis. Curr.Top.Dev.Biol. 71: 131-197.

21. Kimura N, Mizokami, A, Oonuma, T, Sasano, H, and Nagura, H (1993) Immunocytochemical localization of androgen receptor with polyclonal antibody in paraffin-embedded human tissues. J.Histochem.Cytochem. 41: 671-678.

22. Bremner WJ, Millar, MR, Sharpe, RM, and Saunders, PT (1994) Immunohistochemical localization of androgen receptors in the rat testis: evidence for stage-dependent expression and regulation by androgens. Endocrinology 135: 1227-1234.

23. Bellve AR (1993) Purification, culture, and fractionation of spermatogenic cells. Methods Enzymol. 225: 84-113.

24. Hoover F and Goldman, D (1992) Temporally correlated expression of nAChR genes during development of the mammalian retina. Exp.Eye Res. 54: 561-571.

25. Lee S, Lee, DK, Dou, Y, Lee, J, Lee, B et al. (2006) Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. Proc.Natl.Acad.Sci.U.S.A 103: 15392-15397.

26. Hecht NB (1998) Molecular mechanisms of male germ cell differentiation. Bioessays 20: 555-561.

27. Yamamoto M, Takahashi, S, Onodera, K, Muraosa, Y, and Engel, JD (1997) Upstream and downstream of erythroid transcription factor GATA-1. Genes Cells 2: 107-115.

28. Ketola I, Rahman, N, Toppari, J, Bielinska, M, Porter-Tinge, SB et al. (1999) Expression and regulation of transcription factors GATA-4 and GATA-6 in developing mouse testis. Endocrinology 140: 1470-1480.

29. Ketola I, Anttonen, M, Vaskivuo, T, Tapanainen, JS, Toppari, J et al. (2002) Developmental expression and spermatogenic stage specificity of transcription factors GATA-1 and GATA-4 and their cofactors FOG-1 and FOG-2 in the mouse testis. Eur.J.Endocrinol. 147: 397-406.

30. Yomogida K, Ohtani, H, Harigae, H, Ito, E, Nishimune, Y et al. (1994) Developmental stage- and spermatogenic cycle-specific expression of transcription factor GATA-1 in mouse Sertoli cells. Development 120: 1759-1766.

31. Russell LB (1990) Patterns of mutational sensitivity to chemicals in poststem-cell stages of mouse spermatogenesis. Prog.Clin.Biol.Res. 340C: 101-113.

32. Russell LD and Clermont, Y (1977) Degeneration of germ cells in normal, hypophysectomized and hormone treated hypophysectomized rats. Anat.Rec. 187: 347-366.

33. Tan KA, De, GK, Atanassova, N, Walker, M, Sharpe, RM et al. (2005) The role of androgens in sertoli cell proliferation and functional maturation: studies in mice with total or Sertoli cell-selective ablation of the androgen receptor. Endocrinology 146: 2674-2683.

34. Lin RJ, Nagy, L, Inoue, S, Shao, W, Miller, WH, Jr. et al. (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature 391: 811-814.

35. Wong CW and Privalsky, ML (1998) Transcriptional repression by the SMRT-mSin3 corepressor: multiple interactions, multiple mechanisms, and a potential role for TFIIB. Mol.Cell Biol. 18: 5500-5510.

36. Wong JC, Alon, N, Mckerlie, C, Huang, JR, Meyn, MS et al. (2003) Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. Hum.Mol.Genet. 12: 2063-2076.

37. Nadler JJ and Braun, RE (2000) Fanconi anemia complementation group C is required for proliferation of murine primordial germ cells. Genesis. 27: 117-123.

38. Niedzwiedz W, Mosedale, G, Johnson, M, Ong, CY, Pace, P et al. (2004) The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. Mol.Cell 15: 607-620.

39. Cumming RC, Lightfoot, J, Beard, K, Youssoufian, H, O'Brien, PJ et al. (2001) Fanconi anemia group C protein prevents apoptosis in hematopoietic cells through redox regulation of GSTP1. Nat.Med. 7: 814-820.

40. Griswold MD (1998) The central role of Sertoli cells in spermatogenesis. Semin.Cell Dev.Biol. 9: 411-416.

41. De GK, Swinnen, JV, Saunders, PT, Schoonjans, L, Dewerchin, M et al. (2004) A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis. Proc.Natl.Acad.Sci.U.S.A 101: 1327-1332.

42. Holdcraft RW and Braun, RE (2004) Androgen receptor function is required in Sertoli cells for the terminal differentiation of haploid spermatids. Development 131: 459-467.

43. LaVoie HA (2003) The role of GATA in mammalian reproduction. Exp.Biol.Med.(Maywood.) 228: 1282-1290.

44. Viger RS, Taniguchi, H, Robert, NM, and Tremblay, JJ (2004) Role of the GATA family of transcription factors in andrology. J.Androl 25: 441-452.

45. Tsuzuki S and Enver, T (2002) Interactions of GATA-2 with the promyelocytic leukemia zinc finger (PLZF) protein, its homologue FAZF, and the t(11;17)-generated PLZF-retinoic acid receptor alpha oncoprotein. Blood 99: 3404-3410.

46. Merika M and Orkin, SH (1993) DNA-binding specificity of GATA family transcription factors. Mol.Cell Biol. 13: 3999-4010.

47. Ketola I, Anttonen, M, Vaskivuo, T, Tapanainen, JS, Toppari, J et al. (2002) Developmental expression and spermatogenic stage specificity of transcription factors GATA-1 and GATA-4 and their cofactors FOG-1 and FOG-2 in the mouse testis. Eur.J.Endocrinol. 147: 397-406.

48. Watamoto K, Towatari, M, Ozawa, Y, Miyata, Y, Okamoto, M et al. (2003) Altered interaction of HDAC5 with GATA-1 during MEL cell differentiation. Oncogene 22: 9176-9184.

49. Gaughan L, Logan, IR, Cook, S, Neal, DE, and Robson, CN (2002) Tip60 and histone deacetylase 1 regulate androgen receptor activity through changes to the acetylation status of the receptor. J.Biol.Chem. 277: 25904-25913.

50. Qamar I, Park, E, Gong, EY, Lee, HJ, and Lee, K (2009) ARR19 (androgen receptor corepressor of 19 kDa), an antisteroidogenic factor, is regulated by GATA-1 in testicular Leydig cells. J.Biol.Chem. 284: 18021-18032.

51. Jeong BC, Hong, CY, Chattopadhyay, S, Park, JH, Gong, EY et al. (2004) Androgen receptor corepressor-19 kDa (ARR19), a leucine-rich protein that represses the transcriptional activity of androgen receptor through recruitment of histone deacetylase. Mol.Endocrinol. 18: 13-25.

52. Tang CJ, Lin, CY, and Tang, TK (2006) Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. Dev.Biol. 290: 398-410.

53. Shu SK, Liu, Q, Coppola, D, and Cheng, JQ (2010) Phosphorylation and activation of androgen receptor by Aurora-A. J.Biol.Chem. 285: 33045-33053.

54. Khan J, Ezan, F, Cremet, JY, Fautrel, A, Gilot, D et al. (2011) Overexpression of active Aurora-C kinase results in cell transformation and tumour formation. PLoS.One. 6: e26512.

Figure legends

FIG 1. The *Tzfp*$^{GTi/GTi}$ mouse model. (A) Overview of the Tzfp protein with the BTB/POZ and zink finger domains (top), and of the wild-type and mutated *Tzfp* allele upon insertional mutagenesis (middle). The wild-type and mutated *Tzfp* transcript is also included (bottom). (B) Genotyping gel showing fragments indicating samples from wild-type, heterozygous and *Tzfp*$^{GTi/GTi}$ mice. (C) Gene expression analysis of *Mll4* and *Plzf* in wild-type and *Tzfp*$^{GTi/GTi}$ testis reveals no significant difference between wild-type and *Tzfp*$^{GTi/GTi}$ testis. (D) Crosses between heterozygous animals reveal a sex-ratio distortion and fewer than expected homozygous mutant pups. The $\chi 2$ -test was used to determine significance. (E) Weight distribution of testis and epididymus in wild-type and *Tzfp*$^{GTi/GTi}$ mouse reveal a small, but not significant reduction in testis weight in the *Tzfp*$^{GTi/GTi}$ mice.

FIG 2. Tzfp gene expression peaks in pachytene spermatocytes. (A) RT-PCR results showing expression of *Tzfp* in various tissues and organs. (B) Expression of *Tzfp* in testis and meiotic germ cells. (C) A diagram of *Tzfp* expression in developing, juvenile and adult testis with the highest developed germ cell present indicated in the top panel. (D) *In Situ* hybridization with DIG labeled *Tzfp* probe on testis cryo sections. Wild-type (top) and *Tzfp*$^{GTi/GTi}$ testes (bottom) are shown. Left panel shows tubules hybridized with anti-sense probe, indicating positive pachytene spermatocytes in the wild-type (dark color), and absence of signal in the *Tzfp*$^{GTi/GTi}$ sample. Right panel shows tubules hybridized with sense probe, indicating no presence of a sense transcript. The slides are counterstained with haematoxylin (light blue). Scale bar: 20 μm.

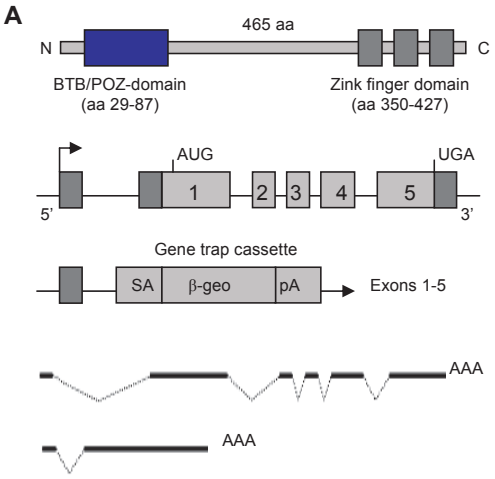
Abbreviations: GC: germ cells, E. /L. pachytene: early and late pachytene spermatocyte, R. /E. spermatids: round and elongated spermatid

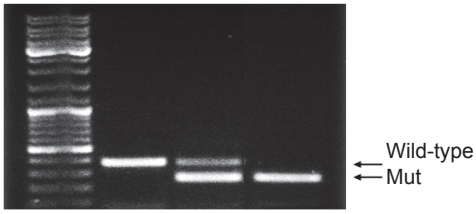
FIG 3. Gene expression analysis. RT-PCR results showing gene expression in wild-type and *Tzfp*$^{GTi/GTi}$ testis at age 12 days and 12 weeks. Transcripts analyzed are *Androgen Receptor*, *Fanconi anemiaC*, *Aie1*, *Gata-1* and *Gata-4*. Statistical analysis was performed by using the two-tailed Student's *t* test: *p < 0.05; **p<0.01; ***p< 0.001.

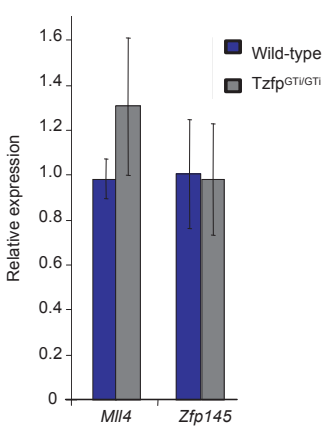
FIG 4. Expression of Androgen Receptor and GATA in wild-type and *Tzfp*$^{GTi/GTi}$ testis. Left panel showing sections of adult wild-type and *Tzfp*$^{GTi/GTi}$ testes incubated with anti-AR (A) and anti-GATA-1 (B) (both displayed in red). Anti-AR samples are co-stained with anti-laminin (green). The slides are counterstained with DAPI (blue). Star (*) in magnified area indicates Sertoli cell. Scale bar: 20 μm. C) Quantification of the antibody signal. Statistical analysis was performed using the two-tailed Student's *t* test: *p < 0.05; **p<0.01; ***p< 0.001. D) Western blot analysis showing the presence of Androgen Receptor, GATA-1 and GATA-4 in adult wild-type and *Tzfp*$^{GTi/GTi}$ testes. β-actin is used as a loading control.

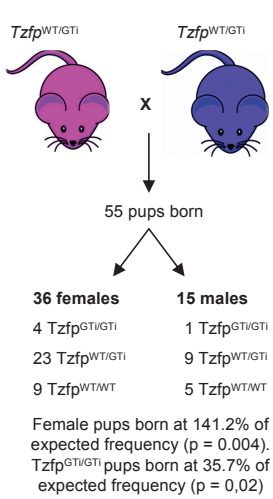
FIG 5. The *Tzfp*$^{GTi/GTi}$ testis develops normally but has a reduced number of apoptotic cells. (A) Stage VII-VIII tubules from wild-type (left) and *Tzfp*$^{GTi/GTi}$ (right) testis. Magnified areas show the presence of spermatogonia (star), pachytene spermatocyte (arrow) and round spermatid (arrowhead). Scale bar: 20 μm. (B) TUNEL assay performed on wild-type (left) and *Tzfp*$^{GTi/GTi}$ (right) testis sections. The histogram indicates quantification of positive cells. Statistical analysis was performed using the two-tailed Student's *t* test: *p < 0.05; **p<0.01; ***p< 0.001.

Figure 1.


A
465 aa
N
C
BTB/POZ-domain
(aa 29-87)
Zink finger domain
(aa 350-427)
AUG
UGA
1
2
3
4
5
5'
3'
Gene trap cassette
SA
β-geo
pA
Exons 1-5
AAA
AAA


B
Wild-type
Mut


C
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
Relative expression
Wild-type
TzfpGTi/GTi
Mll4
Zfp145


D
TzfpWT/GTi
TzfpWT/GTi
X
55 pups born
36 females
15 males
4 TzfpGTi/GTi
1 TzfpGTi/GTi
23 TzfpWT/GTi
9 TzfpWT/GTi
9 TzfpWT/WT
5 TzfpWT/WT
Female pups born at 141.2% of expected frequency (p = 0.004).
TzfpGTi/GTi pups born at 35.7% of expected frequency (p = 0,02)

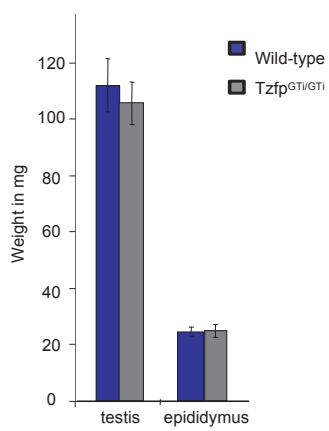

E
120
100
80
60
40
20
0
Weight in mg
Wild-type
TzfpGTi/GTi
testis
epididymus

Figure 2.

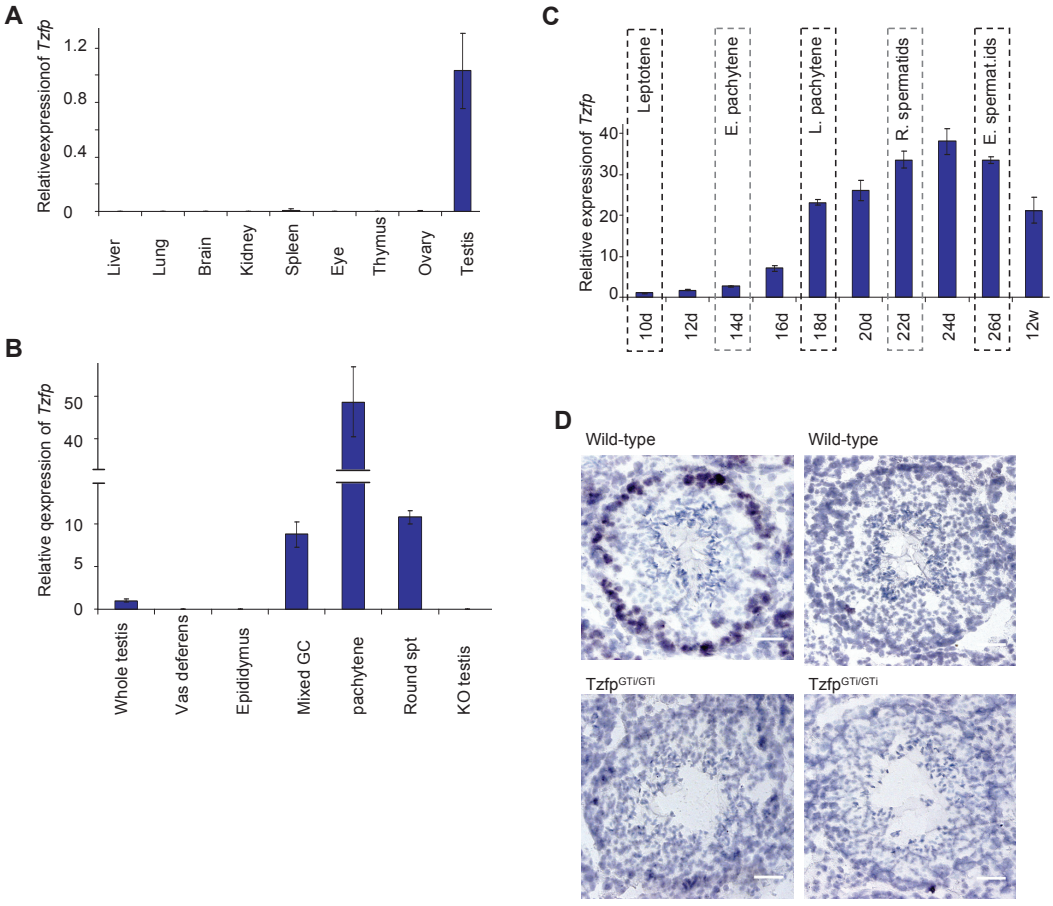

A
Relativeexpressionof Tzfp
1.2
0.8
0.4
0
Liver
Lung
Brain
Kidney
Spleen
Eye
Thymus
Ovary
Testis
B
Relative qexpression of Tzfp
50
40
10
5
0
Whole testis
Vas deferens
Epididymus
Mixed GC
pachytene
Round spt
KO testis
C
Relative expressionof Tzfp
40
30
20
10
0
Leptotene
E. pachytene
L. pachytene
R. spermatids
E. spermat.ids
10d
12d
14d
16d
18d
20d
22d
24d
26d
12w
D
Wild-type
Wild-type
TzfpGTi/GTi
TzfpGTi/GTi

Figure 3.

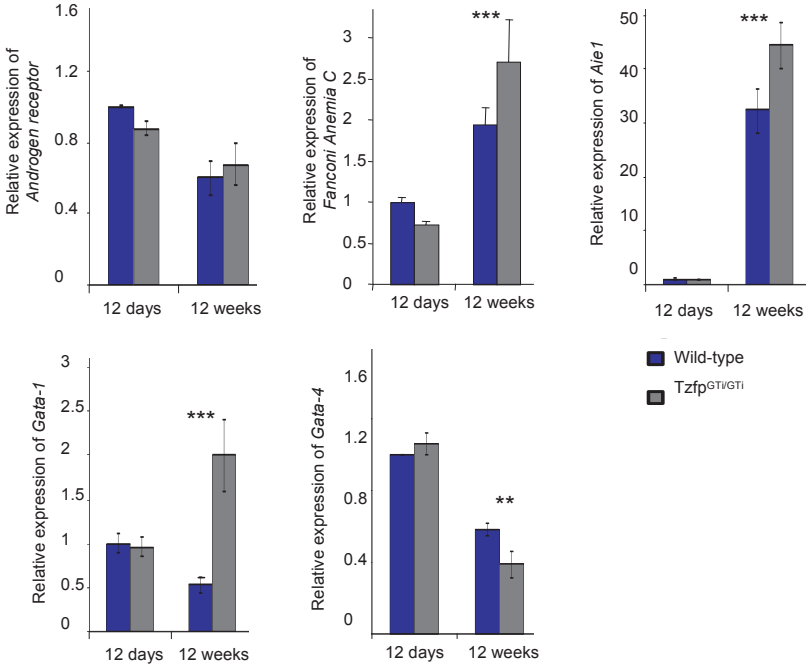

Relative expression of Androgen receptor
1.6
1.2
0.8
0.4
0
12 days
12 weeks
Relative expression of Fanconi Anemia C
***
3
2.5
2
1.5
1
0.5
0
12 days
12 weeks
Relative expression of Aie1
***
50
40
30
20
10
0
12 days
12 weeks
Wild-type
Tzfp GTi/GTi
Relative expression of Gata-1
***
3
2.5
2
1.5
1
0.5
0
12 days
12 weeks
Relative expression of Gata-4
**
1.6
1.2
0.8
0.4
0
12 days
12 weeks

Figure 4.

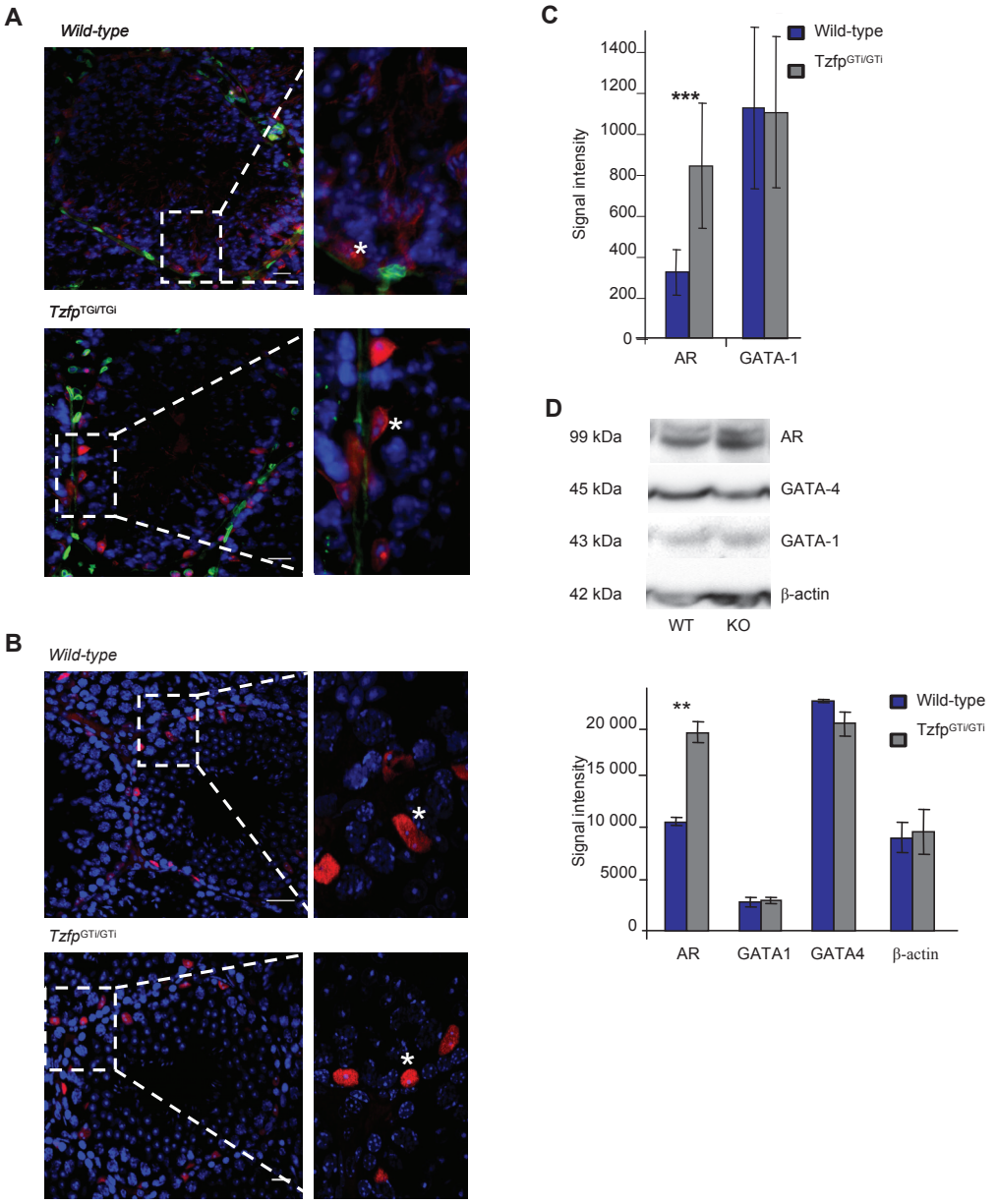
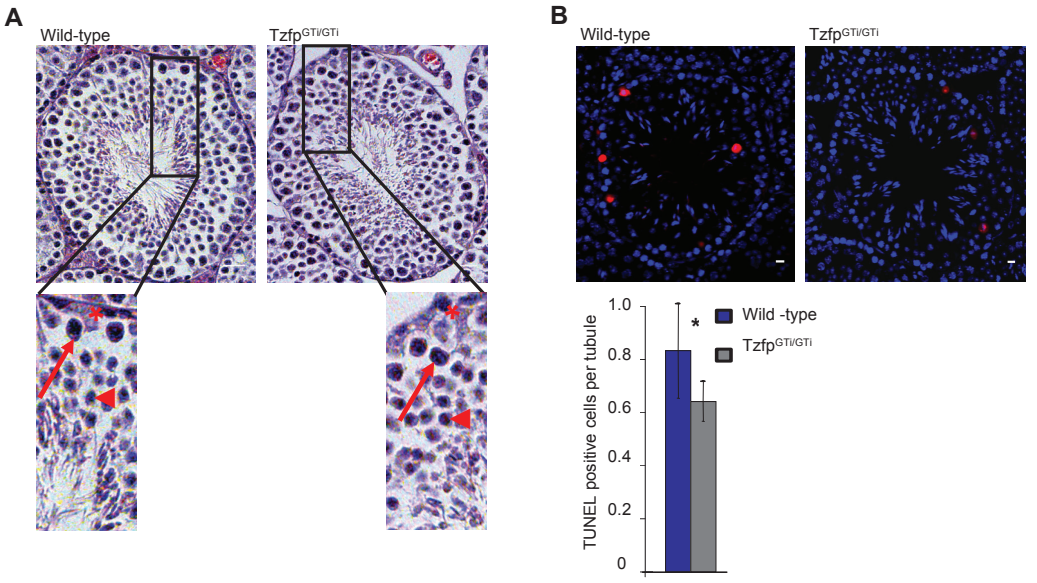

A
Wild-type
TzfpTGi/TGi
B
Wild-type
TzfpGTi/GTi
C
Wild-type
TzfpGTi/GTi
Signal intensity
1400
1200
1000
800
600
400
200
0
***
AR
GATA-1
D
99 kDa
AR
45 kDa
GATA-4
43 kDa
GATA-1
42 kDa
β-actin
WT
KO
Wild-type
TzfpGTi/GTi
Signal intensity
20 000
15 000
10 000
5000
0
**
AR
GATA1
GATA4
β-actin

Figure 5.


A
Wild-type
Tzfp GTi/GTi
B
Wild-type
Tzfp GTi/GTi
TUNEL positive cells per tubule
1.0
0.8
0.6
0.4
0.2
0
*
Wild -type
Tzfp GTi/GTi